

# SEARCH REQUEST FORM

Requestor's Name: \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

## STAFF USE ONLY

Date completed: 08-28-03  
Searcher: Beverly e 4994  
Terminal time: 22  
Elapsed time: \_\_\_\_\_  
CPU time: \_\_\_\_\_  
Total time: 25  
Number of Searches: \_\_\_\_\_  
Number of Databases: 2

Search Site  
\_\_\_\_ STIC  
\_\_\_\_ CM-1  
\_\_\_\_ Pre-S  
Type of Search  
\_\_\_\_ N.A. Sequence  
\_\_\_\_ A.A. Sequence  
\_\_\_\_ Structure  
\_\_\_\_ Bibliographic

Vendors  
\_\_\_\_ IG  
\_\_\_\_ STN  
\_\_\_\_ Dialog  
\_\_\_\_ APS  
\_\_\_\_ Geninfo  
\_\_\_\_ SDC  
\_\_\_\_ DARC/Questel  
\_\_\_\_ Other CGN

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2003, 14:10:52 ; Search time 376 Seconds

(without alignments)  
30.092 Million cell updates/sec

Title: US-09-845-738C-1  
Perfect score: 70  
Sequence: 1 ITRHIMESASL 13

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_AA\_Main:

1: /cgn2\_6/ptodata/1/paa/US06 COMB pep: \*  
2: /cgn2\_6/ptodata/1/paa/US07 COMB pep: \*  
3: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
4: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
5: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
6: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
7: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
8: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
9: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
10: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
11: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
12: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
13: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
14: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
15: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
16: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
17: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
18: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
19: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
20: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
21: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
22: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
23: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
24: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
25: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
26: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
27: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
28: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
29: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
30: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*  
31: /cgn2\_6/ptodata/1/paa/US08 COMB pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	ID	Description
No.						
1	70	100.0	13	23	US-09-845-738A-1	Sequence 1, Appl1
2	70	100.0	14	23	US-09-845-730-1	Sequence 1, Appl1

3	70	100.0	14	23	US-09-846-349A-1	Sequence 1, Appl1
4	70	100.0	15	23	US-09-845-735-1	Sequence 1, Appl1
5	70	100.0	15	23	US-09-845-739-1	Sequence 1, Appl1
6	70	100.0	16	23	US-09-845-730A-1	Sequence 1, Appl1
7	70	100.0	16	23	US-09-846-345-1	Sequence 1, Appl1
8	70	100.0	17	23	US-09-846-345-1	Sequence 1, Appl1
9	70	100.0	17	23	US-09-846-345A-1	Sequence 1, Appl1
10	70	100.0	17	23	US-09-846-346-1	Sequence 1, Appl1
11	70	100.0	18	23	US-09-846-346A-1	Sequence 1, Appl1
12	70	100.0	18	23	US-09-846-346A-1	Sequence 1, Appl1
13	70	100.0	19	21	US-09-724-676-66913	Sequence 66913, A
14	70	100.0	19	21	US-09-724-676-66913	Sequence 66913, A
15	70	100.0	19	21	US-09-724-676-66911	Sequence 66911, A
16	70	100.0	19	21	US-09-724-676-66911	Sequence 66911, A
17	70	100.0	19	21	US-09-724-676-66911	Sequence 66911, A
18	70	100.0	19	21	US-09-724-676-66911	Sequence 66911, A
19	70	100.0	19	21	US-09-724-676-66904	Sequence 66904, A
20	70	100.0	19	21	US-09-724-676-66904	Sequence 66904, A
21	70	100.0	19	21	US-09-724-676-66907	Sequence 66907, A
22	70	100.0	19	21	US-09-724-676-66907	Sequence 66907, A
23	70	100.0	19	21	US-09-724-676-66908	Sequence 66908, A
24	70	100.0	19	21	US-09-724-676-66908	Sequence 66908, A
25	70	100.0	19	21	US-09-724-676-66912	Sequence 66912, A
26	70	100.0	19	21	US-09-724-676-66912	Sequence 66912, A
27	70	100.0	19	21	US-09-724-676-66902	Sequence 66902, A
28	70	100.0	19	21	US-09-724-676-66902	Sequence 66902, A
29	70	100.0	19	21	US-09-724-676-66905	Sequence 66905, A
30	70	100.0	19	21	US-09-724-676-66905	Sequence 66905, A
31	70	100.0	19	21	US-09-724-676-66909	Sequence 66909, A
32	70	100.0	19	21	US-09-724-676-66909	Sequence 66909, A
33	70	100.0	19	21	US-09-724-676-66903	Sequence 66903, A
34	70	100.0	19	21	US-09-724-676-66903	Sequence 66903, A
35	70	100.0	19	21	US-09-724-676-66906	Sequence 66906, A
36	70	100.0	19	21	US-09-724-676-66906	Sequence 66906, A
37	70	100.0	19	21	US-09-724-676-66876	Sequence 66876, A
38	70	100.0	19	21	US-09-724-676-66876	Sequence 66876, A
39	70	100.0	19	21	US-09-724-676-66880	Sequence 66880, A
40	70	100.0	19	21	US-09-724-676-66880	Sequence 66880, A
41	70	100.0	19	21	US-09-724-676-66870	Sequence 66870, A
42	70	100.0	19	21	US-09-724-676-66870	Sequence 66870, A
43	70	100.0	19	21	US-09-724-676-66873	Sequence 66873, A
44	70	100.0	19	21	US-09-724-676-66873	Sequence 66873, A
45	70	100.0	19	21	US-09-724-676-66874	Sequence 66874, A

ALIGNMENTS

RESULT 1  
US-09-845-738A-1  
Sequence 1, Application US/09845738A  
GENERAL INFORMATION:  
APPLICANT: Jackowski, George  
TITLE OF INVENTION: BIOCOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR  
FILE REFERENCE: 2132.040  
CURRENT APPLICATION NUMBER: US/09/845,738A  
CURRENT FILING DATE: 2001-04-30  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 13  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-845-738A-1

Query Match	100.0%	Score 70:	DB 23:	Length 13:
Best Local Similarity	100.0%	Pred. No. 0.00025:		
Matches 13:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
Cy	1	ITRHIMESASL 13		
Db	1	ITRHIMESASL 13		

```

RESULT 2
US-09-845-730-1
; Sequence 1, Application US/09845730
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
; TITLE OF INVENTION: OF 1690 DALTONS
; FILE REFERENCE: 2132.042
; CURRENT APPLICATION NUMBER: US/09/845,730
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-845-730-1

Query Match      100.0%; Score 70; DB 23; Length 14;
Best Local Similarity 100.0%; Pred.No. 0.00027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ITHRIMESASLL 13
        |||||
        2 ITHRIMESASLL 14

Db      1 ITHRIMESASLL 13
        |||||
        1 ITHRIMESASLL 13

RESULT 3
US-09-846-349A-1
; Sequence 1, Application US/09846349A
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; APPLICANT: Vrees, Tammy
; APPLICANT: Yantha, Jason
; APPLICANT: Marshall, John
; APPLICANT: Thatcher, Brad
; TITLE OF INVENTION: Biopolymer Marker Indicative Of Disease State Having A Molecular
; TITLE OF INVENTION: Of 1449 Daltons
; FILE REFERENCE: 2132.034
; CURRENT APPLICATION NUMBER: US/09/846,349A
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-846-349A-1

Query Match      100.0%; Score 70; DB 23; Length 14;
Best Local Similarity 100.0%; Pred.No. 0.00027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ITHRIMESASLL 13
        |||||
        1 ITHRIMESASLL 13

Db      1 ITHRIMESASLL 13
        |||||
        1 ITHRIMESASLL 13

RESULT 4
US-09-845-735-1
; Sequence 1, Application US/09845735
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
; TITLE OF INVENTION: OF 1777 DALTONS
; FILE REFERENCE: 2132.043
; CURRENT APPLICATION NUMBER: US/09/845,735
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1

```

```

; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-845-735-1

Query Match          100.0%; Score 70; DB 23; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ITHRIHWESASILL 13
        |||||
        3 ITHRIHWESASILL 15

Db

RESULT 5
US-09-845-739-1
; Sequence 1, Application US/09845739
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
; FILE REFERENCE: 2132.044
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-845-739-1

Query Match          100.0%; Score 70; DB 23; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ITHRIHWESASILL 13
        |||||
        3 ITHRIHWESASILL 15

Db

RESULT 6
US-09-845-730A-1
; Sequence 1, Application US/09845730A
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
; FILE REFERENCE: 2132.042
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-845-730A-1

Query Match          100.0%; Score 70; DB 23; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ITHRIHWESASILL 13
        |||||
        3 ITHRIHWESASILL 15

Db

RESULT 7
US-09-846-345-1
; Sequence 1, Application US/09846345
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George

```



```

; TITLE OF INVENTION: BIOPOLYMER-MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
; FILE REFERENCE: 2132.045
; CURRENT APPLICATION NUMBER: US/09/846,345
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-846-345-1

Query Match          100.0%; Score 70; DB 23; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHWESASL 13
Db 4 ITHRIHWESASL 16

RESULT 8
US-09-846-344-1
; Sequence 1, Application US/09846344
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
; FILE REFERENCE: 2132.048
; CURRENT APPLICATION NUMBER: US/09/846,344
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-846-344-1

Query Match          100.0%; Score 70; DB 23; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHWESASL 13
Db 4 ITHRIHWESASL 16

RESULT 9
US-09-846-345A-1
; Sequence 1, Application US/09846345A
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
; FILE REFERENCE: 2132.045
; CURRENT APPLICATION NUMBER: US/09/846,345A
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-846-345A-1

Query Match          100.0%; Score 70; DB 23; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHWESASL 13
Db 4 ITHRIHWESASL 16
```

```

Db 4 ITHRIHWESASL 16

RESULT 10
US-09-846-346-1
; Sequence 1, Application US/09846346
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
; FILE REFERENCE: 2132.013
; CURRENT APPLICATION NUMBER: US/09/846,346
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-846-346-1

Query Match          100.0%; Score 70; DB 23; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHWESASL 13
Db 4 ITHRIHWESASL 16

RESULT 11
US-09-846-346A-1
; Sequence 1, Application US/09846346A
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; APPLICANT: Marshall, John
; APPLICANT: Yantha, Jason
; APPLICANT: Vrees, Tammy
; APPLICANT: Thatcher, Brad
; TITLE OF INVENTION: Biopolymer Marker Indicative Of Disease State Having A Molecular
; FILE REFERENCE: 2132.013
; CURRENT APPLICATION NUMBER: US/09/846,346A
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-846-346A-1

Query Match          100.0%; Score 70; DB 23; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHWESASL 13
Db 5 ITHRIHWESASL 17

RESULT 12
US-60-242-679-943
; Sequence 943, Application US/60242679
; GENERAL INFORMATION:
; APPLICANT: Ladunga, Steven Ietvan
; APPLICANT: Spier, Eugene
; APPLICANT: Greenberg, Simon
; APPLICANT: Brandenberger, Ralph
; APPLICANT: Wang, Yu
; APPLICANT: Dubman, Alex
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
```

```
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO00898-PROV
; CURRENT APPLICATION NUMBER: US/60/242,679
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 2265
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 943
; LENGTH: 677
; TYPE: PRT
; ORGANISM: HUMAN
US-60-242-679-943
```

```
Query Match          100.0%; Score 70; DB 31; Length 677;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ITHRIHWSASLL 13
         |||||
Db       321 ITHRIHWSASLL 333
```

```
RESULT 13
US-09-724-676-66913
; Sequence 66913, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66913
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-66913
```

```
Query Match          100.0%; Score 70; DB 21; Length 770;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ITHRIHWSASLL 13
         |||||
Db       471 ITHRIHWSASLL 483
```

```
RESULT 14
US-09-724-676A-66913
; Sequence 66913, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66913
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-66913
```

```
Query Match          100.0%; Score 70; DB 21; Length 770;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ITHRIHWSASLL 13
         |||||
Db       471 ITHRIHWSASLL 483
```

```
RESULT 15
US-09-724-676-66911
; Sequence 66911, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66911
; LENGTH: 785
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-66911
```

```
Query Match          100.0%; Score 70; DB 21; Length 785;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ITHRIHWSASLL 13
         |||||
Db       471 ITHRIHWSASLL 483
```

```
RESULT 16
US-09-724-676A-66911
; Sequence 66911, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66911
; LENGTH: 785
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-66911
```

```
Query Match          100.0%; Score 70; DB 21; Length 785;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ITHRIHWSASLL 13
         |||||
Db       471 ITHRIHWSASLL 483
```

```
RESULT 17
US-09-724-676-66914
; Sequence 66914, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66914
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-66914
```

```
Query Match          100.0%; Score 70; DB 21; Length 793;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 ITHRIHESASLL 13  
| | | | |  
Db 471 ITHRIHESASLL 483

RESULT 18  
US-09-724-676A-66914

; Sequence 66914, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 66914  
; LENGTH: 793  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-66914

Query Match 100.0%; Score 70; DB 21; Length 793;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHESASLL 13  
| | | | |  
Db 471 ITHRIHESASLL 483

RESULT 19  
US-09-724-676-66904

; Sequence 66904, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 66904  
; LENGTH: 804  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-66904

Query Match 100.0%; Score 70; DB 21; Length 804;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHESASLL 13  
| | | | |  
Db 471 ITHRIHESASLL 483

RESULT 20  
US-09-724-676A-66904

; Sequence 66904, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 66904  
; LENGTH: 804  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-66904

Query Match 100.0%; Score 70; DB 21; Length 804;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHESASLL 13  
| | | | |  
Db 471 ITHRIHESASLL 483

RESULT 21  
US-09-724-676-66907

; Sequence 66907, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 66907  
; LENGTH: 812  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-66907

Query Match 100.0%; Score 70; DB 21; Length 812;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHESASLL 13  
| | | | |  
Db 471 ITHRIHESASLL 483

RESULT 22  
US-09-724-676A-66907

; Sequence 66907, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 66907  
; LENGTH: 812  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-66907

Query Match 100.0%; Score 70; DB 21; Length 812;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHESASLL 13  
| | | | |  
Db 471 ITHRIHESASLL 483

RESULT 23  
US-09-724-676-66908

; Sequence 66908, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2

```
; SEQ ID NO 66908
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-66908
```

```
Query Match          100.0%; Score 70; DB 21; Length 822;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ITHRIHWESASLL 13
        |||||
Db       471 ITHRIHWESASLL 483
```

```
RESULT 24
US-09-724-676A-66908
; Sequence 66908, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66908
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-66908
```

```
Query Match          100.0%; Score 70; DB 21; Length 822;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ITHRIHWESASLL 13
        |||||
Db       471 ITHRIHWESASLL 483
```

```
RESULT 25
US-09-724-676-66912
; Sequence 66912, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66912
; LENGTH: 830
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-66912
```

```
Query Match          100.0%; Score 70; DB 21; Length 830;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ITHRIHWESASLL 13
        |||||
Db       471 ITHRIHWESASLL 483
```

```
RESULT 26
US-09-724-676A-66912
; Sequence 66912, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
```

```
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66912
; LENGTH: 830
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-66912
```

```
Query Match          100.0%; Score 70; DB 21; Length 830;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ITHRIHWESASLL 13
        |||||
Db       471 ITHRIHWESASLL 483
```

```
RESULT 27
US-09-724-676-66902
; Sequence 66902, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66902
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-66902
```

```
Query Match          100.0%; Score 70; DB 21; Length 841;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ITHRIHWESASLL 13
        |||||
Db       471 ITHRIHWESASLL 483
```

```
RESULT 28
US-09-724-676A-66902
; Sequence 66902, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66902
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-66902
```

```
Query Match          100.0%; Score 70; DB 21; Length 841;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ITHRIHWESASLL 13
        |||||
Db       471 ITHRIHWESASLL 483
```

```
RESULT 29
```

```
US-09-724-676-66905
; Sequence 66905, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66905
; LENGTH: 849
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-66905

Query Match          100.0%; Score 70; DB 21; Length 849;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 ITHRIHWESASLL 13
Db      471 ITHRIHWESASLL 483

RESULT 30
US-09-724-676A-66905
; Sequence 66905, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66905
; LENGTH: 849
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-66905

Query Match          100.0%; Score 70; DB 21; Length 849;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 ITHRIHWESASLL 13
Db      471 ITHRIHWESASLL 483

RESULT 31
US-09-724-676-66909
; Sequence 66909, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66909
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-66909

Query Match          100.0%; Score 70; DB 21; Length 850;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 ITHRIHWESASLL 13
Db      471 ITHRIHWESASLL 483
```

```
Db      471 ITHRIHWESASLL 483

RESULT 32
US-09-724-676A-66909
; Sequence 66909, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66909
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-66909

Query Match          100.0%; Score 70; DB 21; Length 850;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 ITHRIHWESASLL 13
Db      471 ITHRIHWESASLL 483

RESULT 33
US-09-724-676-66903
; Sequence 66903, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66903
; LENGTH: 869
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-66903

Query Match          100.0%; Score 70; DB 21; Length 869;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 ITHRIHWESASLL 13
Db      471 ITHRIHWESASLL 483

RESULT 34
US-09-724-676A-66903
; Sequence 66903, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66903
; LENGTH: 869
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-66903
```

Query Match 100.0%; Score 70; DB 21; Length 869;  
Best Local Similarity 100.0%; Pred. No. 0.014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHWEASL 13  
Db 471 ITHRIHWEASL 483

RESULT 35  
US-09-724-676-66906  
; Sequence 66906, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 66906  
; LENGTH: 876  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-66906

Query Match 100.0%; Score 70; DB 21; Length 876;  
Best Local Similarity 100.0%; Pred. No. 0.014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHWEASL 13  
Db 471 ITHRIHWEASL 483

RESULT 36  
US-09-724-676A-66906  
; Sequence 66906, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 66906  
; LENGTH: 876  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-66906

Query Match 100.0%; Score 70; DB 21; Length 876;  
Best Local Similarity 100.0%; Pred. No. 0.014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHWEASL 13  
Db 471 ITHRIHWEASL 483

RESULT 37  
US-09-724-676-66876  
; Sequence 66876, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 66876

; LENGTH: 1190  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-66876

Query Match 100.0%; Score 70; DB 21; Length 1190;  
Best Local Similarity 100.0%; Pred. No. 0.019;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHWEASL 13  
Db 876 ITHRIHWEASL 888

RESULT 38  
US-09-724-676A-66876  
; Sequence 66876, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 66876  
; LENGTH: 1190  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-66876

Query Match 100.0%; Score 70; DB 21; Length 1190;  
Best Local Similarity 100.0%; Pred. No. 0.019;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHWEASL 13  
Db 876 ITHRIHWEASL 888

RESULT 39  
US-09-724-676-66880  
; Sequence 66880, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 66880  
; LENGTH: 1198  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-66880

Query Match 100.0%; Score 70; DB 21; Length 1198;  
Best Local Similarity 100.0%; Pred. No. 0.019;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHWEASL 13  
Db 876 ITHRIHWEASL 888

RESULT 40  
US-09-724-676A-66880  
; Sequence 66880, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724, 676A  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 66880  
 ; LENGTH: 1198  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-724-676A-66880

Query Match 100.0%; Score 70; DB 21; Length 1198;  
 Best Local Similarity 100.0%; Pred. No. 0.019;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ITHRIHWESASLT 13  
 |||||  
 Db 876 ITHRIHWESASLT 888

Search completed: August 28, 2003, 14:21:19  
 Job time : 377 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2003, 14:11:12 ; Search time 23 seconds  
(Without alignments)  
16.625 Million cell updates/sec

Title: US-09-845-738C-1  
Perfect score: 70  
Sequence: 1 ITHRHIMESASL 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 140709 seqs, 29413474 residues  
Total number of hits satisfying chosen parameters: 140709

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents AA New:  
1: /cgn2\_6/ptodata/1/paa/PCF\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	100.0	13	US-09-845-738C-1	Sequence 1, Appl
2	70	100.0	705	PCT-US03-06859-4	Sequence 4, Appl
3	70	100.0	705	US-10-379-747-4	Sequence 4, Appl
4	70	100.0	1663	PCT-US03-06859-2	Sequence 2, Appl
5	70	100.0	1663	US-10-379-747-2	Sequence 2, Appl
6	41	58.6	475	US-09-897-516A-4590	Sequence 4590, Ap
7	40	57.1	396	US-10-408-765A-3010	Sequence 3010, Ap
8	38	54.3	460	US-10-613-520-1720	Sequence 1720, Ap
9	38	54.3	586	US-09-319-724B-19	Sequence 19, Appl
10	37.5	53.6	609	US-09-897-516A-7603	Sequence 7603, Ap
11	37	52.9	17	US-09-319-724B-5	Sequence 5, Appl
12	37	52.9	267	US-10-463-190-43	Sequence 43, Appl
13	37	52.9	439	US-09-319-724B-1	Sequence 26, Appl
14	37	52.9	545	US-10-616-263-86	Sequence 14, Appl
15	37	52.9	576	US-09-319-724B-14	Sequence 15, Appl
16	37	52.9	578	US-10-603-108-8858	Sequence 2858, Ap
17	37	52.9	614	US-10-603-108-8858	Sequence 2858, Ap
18	37	52.9	677	US-09-897-516A-5850	Sequence 5850, Ap
19	37	52.9	1210	US-10-624-932-26	Sequence 26, Appl
20	37	52.9	1221	US-10-467-042-4	Sequence 4, Appl
21	36	51.4	150	PCT-US03-14753A-707	Sequence 707, Appl
22	36	51.4	345	US-10-374-780A-2864	Sequence 2864, Ap
23	36	51.4	346	US-10-603-114-4493	Sequence 4493, Ap
24	36	51.4	350	US-10-293-244-1738	Sequence 1738, Ap
25	36	51.4	356	US-10-612-783-5264	Sequence 5264, Ap
26	35.5	50.7	272	US-10-293-244-1486	Sequence 1486, Ap

27	35.5	50.7	358	6	US-10-293-244-1518	Sequence 1518, Ap
28	35.5	50.7	358	6	US-10-408-765A-2327	Sequence 2327, Ap
29	35	50.0	94	7	US-09-485-450-927	Sequence 927, Appl
30	35	50.0	214	6	US-10-603-113-15056	Sequence 15056, A
31	35	50.0	227	6	US-10-603-114-4555	Sequence 4555, Ap
32	35	50.0	337	6	US-10-603-114-6463	Sequence 6463, Ap
33	35	50.0	566	7	US-09-487-610-1550	Sequence 1550, Ap
34	35	50.0	1066	6	US-10-408-765A-1310	Sequence 1310, Ap
35	35	50.0	1211	7	PCT-US02-29560A-205	Sequence 205, Appl
36	35	50.0	1211	7	US-09-487-610-1551	Sequence 1551, Ap
37	34	48.6	210	6	US-10-612-783-6253	Sequence 6253, Ap
38	34	48.6	213	5	US-09-674-546A-778	Sequence 778, Appl
39	34	48.6	566	6	US-10-637-011-21	Sequence 21, Appl
40	34	48.6	694	6	US-10-613-520-797	Sequence 797, Appl
41	34	48.6	694	6	US-10-613-520-1213	Sequence 1213, Ap
42	34	48.6	694	6	US-10-612-783-6891	Sequence 6891, Ap
43	34	48.6	758	6	US-10-341-134-26	Sequence 26, Appl
44	34	48.6	762	6	US-10-341-134-28	Sequence 28, Appl
45	34	48.6	1098	1	PCT-US03-21510-67	Sequence 67, Appl

## ALIGNMENTS

```
RESULT 1
US-09-845-738C-1
; Sequence 1, Application US/09845738C
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
; FILE REFERENCE: 2132 040
; CURRENT APPLICATION NUMBER: US/09/845, 738C
; CURRENT FILING DATE: 2001-04-30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-845-738C-1

Query Match      100.0%; Score 70; DB 5; Length 13;
Best Local Similarity 100.0%; Pred. No. 66-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ITHRHIMESASL 13
Db      1 ITHRHIMESASL 13

RESULT 2
PCT-US03-06859-4
; Sequence 4, Application PC/TUS0306859
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568B-061
; CURRENT APPLICATION NUMBER: PCT/US03/06859
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/379,747
; PRIOR FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: 60/365,034
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/366,420
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 60/365,477
; PRIOR FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 4
; LENGTH: 705
; TYPE: PRT
```

ORGANISM: Homo sapiens  
PCT-US03-06859-4

Query Match 100.0%; Score 70; DB 1; Length 705;  
Best Local Similarity 100.0%; Pred. No. 0.00033;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ITHRIHWESASLL 13  
|||||  
Db 349 ITHRIHWESASLL 361

## RESULT 3

US-10-379-747-4  
Sequence 4, Application US/10379747  
GENERAL INFORMATION:  
APPLICANT: Burgess, Catherine E.;  
APPLICANT: Chant, John S.;  
APPLICANT: Chaudhuri, Amitabha ;  
APPLICANT: Edinger, Shlomit R.;  
APPLICANT: Gangolli, Esna A.;  
APPLICANT: Malvankar, Uriel M.;  
APPLICANT: Miller, Charles E.;  
APPLICANT: Ooi, Chean Eng;  
APPLICANT: Paturajan, Meera ;  
APPLICANT: Rastelli, Luca ;  
APPLICANT: Rieger, Daniel K.;  
APPLICANT: Shimkets, Richard A.;  
APPLICANT: Zernusen, Bryan D.  
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
FILE REFERENCE: 21402-568B  
CURRENT APPLICATION NUMBER: US/10/379,747  
CURRENT FILING DATE: 2003-03-05  
PRIOR APPLICATION NUMBER: 60/365,034  
PRIOR FILING DATE: 2002-03-15  
PRIOR APPLICATION NUMBER: 60/366,420  
PRIOR FILING DATE: 2002-03-21  
PRIOR APPLICATION NUMBER: 60/365,477  
PRIOR FILING DATE: 2002-03-19  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: Cursseqdist version 0.1  
SEQ ID NO 4  
LENGTH: 705  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-379-747-4

Query Match 100.0%; Score 70; DB 6; Length 705;  
Best Local Similarity 100.0%; Pred. No. 0.00033;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ITHRIHWESASLL 13  
|||||  
Db 349 ITHRIHWESASLL 361

## RESULT 4

PCT-US03-06859-2  
Sequence 2, Application PC/TUS0306859  
GENERAL INFORMATION:  
APPLICANT: Curagen Corporation, et al  
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
FILE REFERENCE: 21402-568B-061  
CURRENT APPLICATION NUMBER: PCT/US03/06859  
CURRENT FILING DATE: 2003-03-06  
PRIOR APPLICATION NUMBER: 10/379,747  
PRIOR FILING DATE: 2003-03-05  
PRIOR APPLICATION NUMBER: 60/365,034  
PRIOR FILING DATE: 2002-03-15  
PRIOR APPLICATION NUMBER: 60/366,420  
PRIOR FILING DATE: 2002-03-21  
PRIOR APPLICATION NUMBER: 60/365,477

PRIOR FILING DATE: 2002-03-19  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: Cursseqdist version 0.1  
SEQ ID NO 2  
LENGTH: 1663  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US03-06859-2

Query Match 100.0%; Score 70; DB 1; Length 1663;  
Best Local Similarity 100.0%; Pred. No. 0.00078;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ITHRIHWESASLL 13  
|||||  
Db 1307 ITHRIHWESASLL 1319

## RESULT 5

US-10-379-747-2  
Sequence 2, Application US/10379747  
GENERAL INFORMATION:  
APPLICANT: Burgess, Catherine E.;  
APPLICANT: Chant, John S.;  
APPLICANT: Chaudhuri, Amitabha ;  
APPLICANT: Edinger, Shlomit R.;  
APPLICANT: Gangolli, Esna A.;  
APPLICANT: Malvankar, Uriel M.;  
APPLICANT: Miller, Charles E.;  
APPLICANT: Ooi, Chean Eng;  
APPLICANT: Paturajan, Meera ;  
APPLICANT: Rastelli, Luca ;  
APPLICANT: Rieger, Daniel K.;  
APPLICANT: Shimkets, Richard A.;  
APPLICANT: Zernusen, Bryan D.  
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
FILE REFERENCE: 21402-568B  
CURRENT APPLICATION NUMBER: US/10/379,747  
CURRENT FILING DATE: 2003-03-05  
PRIOR APPLICATION NUMBER: 60/365,034  
PRIOR FILING DATE: 2002-03-15  
PRIOR APPLICATION NUMBER: 60/366,420  
PRIOR FILING DATE: 2002-03-21  
PRIOR APPLICATION NUMBER: 60/365,477  
PRIOR FILING DATE: 2002-03-19  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: Cursseqdist version 0.1  
SEQ ID NO 2  
LENGTH: 1663  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-379-747-2

Query Match 100.0%; Score 70; DB 6; Length 1663;  
Best Local Similarity 100.0%; Pred. No. 0.00078;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ITHRIHWESASLL 13  
|||||  
Db 1307 ITHRIHWESASLL 1319

## RESULT 6

US-09-897-516A-4590  
Sequence 4590, Application US/09897516A  
GENERAL INFORMATION:  
APPLICANT: Corbin, David R.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Hueising, Joseph E.  
APPLICANT: Malvar, Thomas W.  
APPLICANT: Krasomil-Osterfeld, Karina C.

```

; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 4590
; LENGTH: 475
; TYPE: PR1
; ORGANISM: Xenorhabdus sp.
US-09-897-516A-4590

```

```

Query Match
Best Local Similarity 58.6%; Score 41; DB 5; Length 475;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1 THRIHESASL 12
Db 183 LHRHESREL 194

```

```

RESULT 7
US-10-408-765A-3010
; Sequence 3010, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Watnick, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3010:
; LENGTH: 396
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-408-765A-3010

```

```

Query Match
Best Local Similarity 57.1%; Score 40; DB 6; Length 396;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 3 HRIHESASL 13
Db 175 HSHWQAKSL 185

```

```

RESULT 8
US-10-613-520-1720
; Sequence 1720, Application US/10613520
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; APPLICANT: Wu, Wei
; APPLICANT: Maoloug, Lu
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)E
; CURRENT APPLICATION NUMBER: US/10/613,520
; CURRENT FILING DATE: 2003-07-02
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 10/389,566
; PRIOR FILING DATE: 2003-03-14

```

```

; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2578
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1720
; LENGTH: 460
; TYPE: PR1
; ORGANISM: Burkholderia fungorum
US-10-613-520-1720

```

```

Query Match
Best Local Similarity 54.3%; Score 38; DB 6; Length 460;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 3 HRIHESASL 13
Db 443 HDVHMSSEPL 453

```

```

RESULT 9
US-09-319-724B-19
; Sequence 19, Application US/09319724B
; GENERAL INFORMATION:
; APPLICANT: VERIGEN
; APPLICANT: LENZEN, Gerlinda
; APPLICANT: STROBERG, Arthur Donny
; APPLICANT: SUGASAWA, Toshinari
; APPLICANT: MOROOKA, Shigeaki
; TITLE OF INVENTION: MAMMALIAN ICYP (IODOCYANOPINDOLOL) RECEPTOR AND ITS APPLICATIONS
; FILE REFERENCE: 53356-5001-US
; CURRENT APPLICATION NUMBER: US/09/319,724B
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: EP 96402719.7
; PRIOR FILING DATE: 1996-12-12
; PRIOR APPLICATION NUMBER: PCT/EP97/07339
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 19
; LENGTH: 586
; TYPE: PR1
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Genbank No. D87444
US-09-319-724B-19

```

```

Query Match
Best Local Similarity 54.3%; Score 38; DB 5; Length 586;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 THRIHES 9
Db 216 SHRIHES 223

```

```

RESULT 10
US-09-897-516A-7603
; Sequence 7603, Application US/09897516A
; GENERAL INFORMATION:
; APPLICANT: Cordin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A

```

```
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 7603
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516A-7603
```

```
Query Match          53.6%; Score 37.5; DB 5; Length 609;
Best Local Similarity 50.0%; Pred. No. 75;
Matches 8; Conservative 1; Mismatches 4; Indels 3; Gaps 1;
```

```
Oy      1 ITHRIHWE--SASLL 13
         |||||
Db      128 IAHVHWEEQOGGSLL 143
```

```
RESULT 11
US-09-319-724B-5
; Sequence 5, Application US/09319724B
; GENERAL INFORMATION:
; APPLICANT: LENZEN, Gerlinde
; APPLICANT: VERITIGEN
; APPLICANT: STROSBERG, Arthur Donny
; APPLICANT: SUGASAWA, Toshinari
; APPLICANT: MOROOKA, Shigeaki
; TITLE OF INVENTION: MAMMALIAN ICYP (IODOCYANOPINDOLOL) RECEPTOR AND ITS APPLICATIONS
; FILE REFERENCE: 53356-5001-US
; CURRENT APPLICATION NUMBER: US/09/319,724B
; PRIOR FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: EP 96402719.7
; PRIOR FILING DATE: 1996-12-12
; PRIOR APPLICATION NUMBER: PCT/EP97/07339
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; TYPE: PRT
; LENGTH: 17
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Fragment generated by acidic cleavage of polypeptide able to bind
US-09-319-724B-5
```

```
Query Match          52.9%; Score 37; DB 5; Length 17;
Best Local Similarity 85.7%; Pred. No. 2.5;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      3 HRIHWES 9
         |||||
Db      7 HRIHWFS 13
```

```
RESULT 12
US-10-463-190-43
; Sequence 43, Application US/10463190
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: INCREASING BONE MINERALIZATION
; FILE REFERENCE: 240083.508C2
; CURRENT APPLICATION NUMBER: US/10/463,190
; CURRENT FILING DATE: 2003-06-16
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 43
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-463-190-43
```

```
Query Match          52.9%; Score 37; DB 6; Length 267;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      2 THRIHWES 9
         :|||:
Db      154 SHEVHWET 161
```

```
RESULT 13
US-09-319-724B-1
; Sequence 1, Application US/09319724B
; GENERAL INFORMATION:
; APPLICANT: VERITIGEN
; APPLICANT: LENZEN, Gerlinde
; APPLICANT: STROSBERG, Arthur Donny
; APPLICANT: SUGASAWA, Toshinari
; APPLICANT: MOROOKA, Shigeaki
; TITLE OF INVENTION: MAMMALIAN ICYP (IODOCYANOPINDOLOL) RECEPTOR AND ITS APPLICATIONS
; FILE REFERENCE: 53356-5001-US
; CURRENT APPLICATION NUMBER: US/09/319,724B
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: EP 96402719.7
; PRIOR FILING DATE: 1996-12-12
; PRIOR APPLICATION NUMBER: PCT/EP97/07339
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-319-724B-1
```

```
Query Match          52.9%; Score 37; DB 5; Length 439;
Best Local Similarity 85.7%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      3 HRIHWES 9
         |||||
Db      88 HRIHWFS 94
```

```
RESULT 14
US-10-616-263-26
; Sequence 26, Application US/10616263
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Werberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES ENCODING THEM
; FILE REFERENCE: 00766.00103.5
; CURRENT APPLICATION NUMBER: US/10/616,263
; CURRENT FILING DATE: 2003-07-08
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
```

SEQ ID NO 26  
LENGTH: 545  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-616-263-26

Query Match  
Best Local Similarity 52.9%; Score 37; DB 6; Length 545;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 HRIHWS 9  
Db 175 HRIHWFS 181

RESULT 15  
US-09-319-724B-14  
Sequence 14, Application US/09319724B  
GENERAL INFORMATION:  
APPLICANT: VETIGEN  
APPLICANT: LENZEN, Gerlinde  
APPLICANT: STROSEBERG, Arthur Donny  
APPLICANT: SUGASAWA, Toshinari  
APPLICANT: MOROOKA, Shigeaki  
TITLE OF INVENTION: MAMMALIAN ICYP (IODOCYANOPINDOL) RECEPTOR AND ITS APPLICATIONS  
FILE REFERENCE: 53356-5001-US  
CURRENT APPLICATION NUMBER: US/09/319,724B  
CURRENT FILING DATE: 2003-07-24  
PRIOR APPLICATION NUMBER: EP 96402719.7  
PRIOR FILING DATE: 1996-12-12  
PRIOR APPLICATION NUMBER: PCT/EP97/07339  
PRIOR FILING DATE: 1997-12-12  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 14  
LENGTH: 576  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-319-724B-14

Query Match  
Best Local Similarity 52.9%; Score 37; DB 5; Length 576;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 HRIHWS 9  
Db 206 HRIHWFS 212

RESULT 16  
US-09-319-724B-15  
Sequence 15, Application US/09319724B  
GENERAL INFORMATION:  
APPLICANT: VETIGEN  
APPLICANT: LENZEN, Gerlinde  
APPLICANT: STROSEBERG, Arthur Donny  
APPLICANT: SUGASAWA, Toshinari  
APPLICANT: MOROOKA, Shigeaki  
TITLE OF INVENTION: MAMMALIAN ICYP (IODOCYANOPINDOL) RECEPTOR AND ITS APPLICATIONS  
FILE REFERENCE: 53356-5001-US  
CURRENT APPLICATION NUMBER: US/09/319,724B  
CURRENT FILING DATE: 2003-07-24  
PRIOR APPLICATION NUMBER: EP 96402719.7  
PRIOR FILING DATE: 1996-12-12  
PRIOR APPLICATION NUMBER: PCT/EP97/07339  
PRIOR FILING DATE: 1997-12-12  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 15  
LENGTH: 578  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:

NAME/KEY: MISC FEATURE  
OTHER INFORMATION: SM binding protein  
US-09-319-724B-15

Query Match  
Best Local Similarity 52.9%; Score 37; DB 5; Length 578;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 HRIHWS 9  
Db 208 HRIHWFS 214

RESULT 17  
US-10-603-108-2858  
Sequence 2858, Application US/10603108  
GENERAL INFORMATION:  
APPLICANT: Gary L. Bretton  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS  
FILE REFERENCE: PAT03-14  
CURRENT APPLICATION NUMBER: US/10/603,108  
CURRENT FILING DATE: 2003-06-24  
PRIOR APPLICATION NUMBER: US 09/540,263  
PRIOR FILING DATE: 2000-04-04  
PRIOR APPLICATION NUMBER: US 60/125,416  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 3840  
SEQ ID NO 2858  
LENGTH: 614  
TYPE: PRT  
ORGANISM: M. catarrhalis  
US-10-603-108-2858

Query Match  
Best Local Similarity 52.9%; Score 37; DB 6; Length 614;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 ITHRIHWSASL 12  
Db 316 LTHWVLMETAGL 327

RESULT 18  
US-09-897-516A-5850  
Sequence 5850, Application US/09897516A  
GENERAL INFORMATION:  
APPLICANT: Corbin, David R.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Huesing, Joseph E.  
APPLICANT: Malvar, Thomas M.  
APPLICANT: Krasomil-Osterfeld, Karina C.  
APPLICANT: Slater, Steven C.  
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-21(51847)B  
CURRENT APPLICATION NUMBER: US/09/897,516A  
CURRENT FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US 60/215,161  
PRIOR FILING DATE: 2000-06-30  
NUMBER OF SEQ ID NOS: 8415  
SEQ ID NO 5850  
LENGTH: 677  
TYPE: PRT  
ORGANISM: Xenorhabdus sp.  
US-09-897-516A-5850

Query Match  
Best Local Similarity 52.9%; Score 37; DB 5; Length 677;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 IHWSASL 12

Db 587 IHWSAEV 594

RESULT 19  
US-10-624-932-26

; Sequence 26, Application US/10624932  
; GENERAL INFORMATION:  
; APPLICANT: Taudier, Raymond  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Raestelli, Luca  
; APPLICANT: Spaderna, Steven  
; APPLICANT: Shinkels, Richard  
; APPLICANT: Zernusen, Bryan  
; APPLICANT: Spylek, Kimberly  
; APPLICANT: Shenoy, Suresh  
; APPLICANT: Li, Li  
; APPLICANT: Gusev, Vladimir  
; APPLICANT: Grosse, William  
; APPLICANT: Alsbrook, John  
; APPLICANT: Lepley, Denise  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Ellerman, Karen  
; APPLICANT: MacDougall, John  
; APPLICANT: Stone, David  
; APPLICANT: Smithson, Glenda  
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-074 US  
; CURRENT APPLICATION NUMBER: US/10/624,932  
; CURRENT FILING DATE: 2003-07-21  
; PRIOR APPLICATION NUMBER: 09/918,779  
; PRIOR FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: 60/221,409  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: 60/222,840  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,752  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/223,762  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/223,770  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/223,769  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/225,146  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/225,392  
; PRIOR FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: 60/225,470  
; PRIOR FILING DATE: 2000-08-15  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 26  
; LENGTH: 1210  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1185)  
; OTHER INFORMATION: wherein Xaa is any amino acid as defined in the  
; OTHER INFORMATION: specification  
US-10-624-932-26

Query Match 52.9%; Score 37; DB 6; Length 1210;  
Best Local Similarity 54.5%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 HRIHWSASL 13  
DB 428 HRFHWSRSL 438

RESULT 20  
US-10-467-042-4

; Sequence 4, Application US/10467042  
; GENERAL INFORMATION:  
; APPLICANT: WARREN, Bridget A.; HONCHELL, Cynthia D.;  
; APPLICANT: LU, Yan; CHAWLA, Nandinder K.;  
; APPLICANT: BURFORD, Neil; DELBEANE, Angelo M.;  
; APPLICANT: GANDHI, Ameena R.; BAUGHN, Mariah R.;  
; APPLICANT: GRIFFIN, Jennifer A.; GIETZEN, Kimberly J.;  
; APPLICANT: LU, Dyrng Aina M.; ISON, Craig H.;  
; APPLICANT: RAMKUMAR, Jayalaxmi; TANG, Y. Tom;  
; APPLICANT: LAL, Preeti G.; BOROWSKY, Mark L.;  
; APPLICANT: DUGGAN, Brendan M.; HAPALIA, April J.A.;  
; APPLICANT: ARVIZU, Chandra S.; THANCAVELU, Kavitha;  
; APPLICANT: YAO, Monique G.; ELLIOTT, Vicki S.;  
; APPLICANT: DING, Li; YUE, Henry;  
; APPLICANT: LEE, Sally; SMARNAKAR, Anita;  
; APPLICANT: TRAN, Uyen K.; XU, Yuming  
; TITLE OF INVENTION: PROTEIN MODIFICATION AND MAINTENANCE MOLECULES  
; FILE REFERENCE: PI-0361 USN  
; CURRENT APPLICATION NUMBER: US/10/467,042  
; CURRENT FILING DATE: 2003-07-31  
; PRIOR APPLICATION NUMBER: PCT/US02/02813  
; PRIOR FILING DATE: 2002-01-30  
; PRIOR APPLICATION NUMBER: US 60/265,705  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: US 60/266,762  
; PRIOR FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: US 60/269,581  
; PRIOR FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/271,198  
; PRIOR FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: US 60/272,813  
; PRIOR FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: US 60/275,586  
; PRIOR FILING DATE: 2001-03-13  
; PRIOR APPLICATION NUMBER: US 60/278,505  
; PRIOR FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: US 60/280,539  
; PRIOR FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PERL Program  
; SEQ ID NO 4  
; LENGTH: 1221  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: 1719204CD1  
US-10-467-042-4

Query Match 52.9%; Score 37; DB 6; Length 1221;  
Best Local Similarity 54.5%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 HRIHWSASL 13  
DB 434 HRFHWSRSL 444

RESULT 21  
PCT-US02-14753A-707

; Sequence 707, Application PC/TUS0214753A  
; GENERAL INFORMATION:  
; APPLICANT: Corixa Corporation  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, David C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.

APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Derrick  
APPLICANT: Li, Samuel X.  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A. W.  
APPLICANT: Hepler, William T.  
APPLICANT: Hurai, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
APPLICANT: Vinals y de Basols, Carlot  
APPLICANT: Foy, Teresa M.  
APPLICANT: Matanabe, Yoshihiro  
APPLICANT: Deng, Ta  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.42725PC  
CURRENT APPLICATION NUMBER: PCT/US02/14753A  
CURRENT FILING DATE: 2002-05-09  
NUMBER OF SEQ ID NOS: 1033  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 707  
LENGTH: 150  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US02-14753A-707

Query Match 51.4%; Score 36; DB 1; Length 150;  
Best Local Similarity 50.0%; Pred. No. 33;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ITHRIHWSA 10  
Db 133 LAHRHRMNA 142

RESULT 22  
US-10-374-780A-2864  
Sequence 2864, Application US/10374780A  
GENERAL INFORMATION:  
APPLICANT: Sherman, Bradley K  
APPLICANT: Riechmann, Jose Luis  
APPLICANT: Jiang, Cai-Zhong  
APPLICANT: Heard, Jacqueline B  
APPLICANT: Haake, Volker  
APPLICANT: Creelman, Robert A  
APPLICANT: Ratcliffe, Oliver  
APPLICANT: Adam, Luc J  
APPLICANT: Reuber, T. Lynne  
APPLICANT: Keddie, James  
APPLICANT: Broun, Pierre E  
APPLICANT: Pilgrim, Marsha L  
APPLICANT: Dubell III, Arnold T  
APPLICANT: Pineda, Omaira  
APPLICANT: Yu, Guo-Liang  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS  
FILE REFERENCE: MBI-0047 CIP  
CURRENT APPLICATION NUMBER: US/10/374,780A  
CURRENT FILING DATE: 2003-02-25  
PRIOR APPLICATION NUMBER: 09/837,944  
PRIOR FILING DATE: 2001-04-18  
PRIOR APPLICATION NUMBER: 60/310,847  
PRIOR FILING DATE: 2001-08-09  
PRIOR APPLICATION NUMBER: 09/934,455  
PRIOR FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/336,049  
PRIOR FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/338,692  
PRIOR FILING DATE: 2001-12-11  
PRIOR APPLICATION NUMBER: 10/171,468

PRIOR FILING DATE: 2002-06-14  
PRIOR APPLICATION NUMBER: 10/225,066  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: 10/225,067  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: 10/225,068  
PRIOR FILING DATE: 2002-08-09  
NUMBER OF SEQ ID NOS: 2906  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 2864  
LENGTH: 345  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
OTHER INFORMATION: G2426  
US-10-374-780A-2864

Query Match 51.4%; Score 36; DB 6; Length 345;  
Best Local Similarity 50.0%; Pred. No. 75;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ITHRIHWSASL 12  
Db 146 ISHMAQWESARL 157

RESULT 23  
US-10-603-114-4493  
Sequence 4493, Application US/10603114  
GENERAL INFORMATION:  
APPLICANT: GARY BRETTON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
FILE REFERENCE: 2709.1002-001  
CURRENT APPLICATION NUMBER: US/10/603,114  
CURRENT FILING DATE: 2003-06-24  
PRIOR APPLICATION NUMBER: US/09/543,681  
PRIOR FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 4493  
LENGTH: 346  
TYPE: PRT  
ORGANISM: Proteus mirabilis  
US-10-603-114-4493

Query Match 51.4%; Score 36; DB 6; Length 346;  
Best Local Similarity 33.3%; Pred. No. 75;  
Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 2 THRIHWSASL 13  
Db 121 THKVFMDNATIV 132

RESULT 24  
US-10-293-244-1738  
Sequence 1738, Application US/10293244  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc.  
APPLICANT: Yang, Y. Tom et al  
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides  
FILE REFERENCE: 21272-029  
CURRENT APPLICATION NUMBER: US/10/293,244  
CURRENT FILING DATE: 2002-11-12  
PRIOR APPLICATION NUMBER: Not Yet Assigned  
PRIOR FILING DATE: 2003-01-30  
PRIOR APPLICATION NUMBER: 09/728,422  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: 09/693,325  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 09/663,561

```
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO: 1738
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-244-1738
```

```
Query Match          51.4%; Score 36; DB 6; Length 350;
Best Local Similarity 55.6%; Pred. No. 76;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 ITHRIHWES 9
       :|||:|:|:|:|
Db      315 VTHNIHYET 323
```

```
RESULT 25
US-10-612-783-5264
; Sequence 5264, Application US/10612783
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(53373)A
; CURRENT APPLICATION NUMBER: US/10/612,783
; CURRENT FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 7098
; SEQ ID NO: 5264
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4577_184395C.1.pep
US-10-612-783-5264
```

```
Query Match          51.4%; Score 36; DB 6; Length 356;
Best Local Similarity 50.0%; Pred. No. 78;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
OY      1 ITHRIHWESASL 12
       :|||:|:|:|:|
Db      148 LSHMAQWESARL 159
```

```
RESULT 26
US-10-293-244-3486
; Sequence 3486, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hysed, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
```

```
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO: 3486
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-244-3486
```

```
Query Match          50.7%; Score 35.5; DB 6; Length 272;
Best Local Similarity 47.1%; Pred. No. 72;
Matches 8; Conservative 1; Mismatches 3; Indels 5; Gaps 1;
```

```
OY      2 THR-----IHWESASL 13
       :|||:|:|:|:|
Db      169 SHRPVGMDIHWKYSKL 185
```

```
RESULT 27
US-10-293-244-1518
; Sequence 1518, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hysed, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO: 1518
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-244-1518
```

```
Query Match          50.7%; Score 35.5; DB 6; Length 358;
Best Local Similarity 47.1%; Pred. No. 95;
Matches 8; Conservative 1; Mismatches 3; Indels 5; Gaps 1;
```

```
OY      2 THR-----IHWESASL 13
       :|||:|:|:|:|
```



```

Db          255 SHRPVGMIDHWEKVS.KL 271

RESULT 28
US-10-408-765A-2327
; Sequence 2327, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Baby, Soin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Marnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2327
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2327

Query Match          50.7%; Score 35.5; DB 6; Length 358;
Best Local Similarity 47.1%; Pred. No. 95;
Matches 8; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

QY          2 THR-----IHVESASLL 13
           :|||:|||||
Db          255 SHRPVGMIDHWEKVS.KL 271

RESULT 29
US-60-485-450-927
; Sequence 927, Application US/60485450
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERPERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; FILE REFERENCE: CLO01470
; CURRENT APPLICATION NUMBER: US/60/485,450
; CURRENT FILING DATE: 2003-07-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 927
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-485-450-927

Query Match          50.0%; Score 35; DB 7; Length 94;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY          2 THRIHW 7
           :|||:|
Db          70 SHRLHW 75

RESULT 30
US-10-603-113-15056
; Sequence 15056, Application US/10603113
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

```

```

FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/10/603.113
CURRENT FILING DATE: 2003-06-24
PRIOR APPLICATION NUMBER: US/09/248.796
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 28206
SEQ ID NO: 15056
LENGTH: 214
TYPE: PRT
ORGANISM: Candida albicans
US-10-603-113-15056

Query Match      50.0%; Score 35; DB 6; Length 214;
Best Local Similarity 71.4%; Pred. No. 68;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      4 RHMESA 10
      |||||:
Db      41 RHIMENS 47

RESULT 31
US-10-603-114-4555
Sequence 4555, Application US/10603114
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/10/603.114
CURRENT FILING DATE: 2003-06-24
PRIOR APPLICATION NUMBER: US/09/543.681
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO: 4555
LENGTH: 227
TYPE: PRT
ORGANISM: Proteus mirabilis
US-10-603-114-4555

Query Match      50.0%; Score 35; DB 6; Length 227;
Best Local Similarity 50.0%; Pred. No. 72;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      3 HRIHESASL 12
      |||||:
Db      47 HRSHWKGDSI 56

RESULT 32
US-10-603-114-6463
Sequence 6463, Application US/10603114
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/10/603.114
CURRENT FILING DATE: 2003-06-24
PRIOR APPLICATION NUMBER: US/09/543.681
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO: 6463
LENGTH: 337
TYPE: PRT
ORGANISM: Proteus mirabilis
US-10-603-114-6463

Query Match      50.0%; Score 35; DB 6; Length 337;

```

Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 3 HRIHESAS 10  
Db 113 HRLFWDSA 120

RESULT 33  
US-60-487-610-1550  
; Sequence 1550, Application US/60487610  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: HUANG, Hongjin  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,  
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01469  
; CURRENT APPLICATION NUMBER: US/60/487,610  
; CURRENT FILING DATE: 2003-07-17  
; NUMBER OF SEQ ID NOS: 97101  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1550  
; LENGTH: 566  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-487-610-1550

Query Match 50.0%; Score 35; DB 7; Length 566;  
Best Local Similarity 55.6%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 3 HRIHESAS 11  
Db 444 HRFHWSRCS 452

RESULT 34  
US-10-408-765A-1310  
; Sequence 1310, Application US/10408765A  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Boia D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary W.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1310  
; LENGTH: 1066  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-408-765A-1310

Query Match 50.0%; Score 35; DB 6; Length 1086;  
Best Local Similarity 62.5%; Pred. No. 3.5e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ITHRIHME 8  
Db 84 LHLHLHME 91

RESULT 35  
PCT-US02-29560A-205  
; Sequence 205, Application PC/TUS0229560A

; GENERAL INFORMATION:  
; APPLICANT: Afari, Daniel  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Hevezl, Peter A.  
; APPLICANT: Mack, David H.  
; APPLICANT: Wilson, Keith E.  
; APPLICANT: Zlotnik, Albert  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer  
; FILE REFERENCE: 018501-002710PC  
; CURRENT APPLICATION NUMBER: PCT/US02/29560A  
; CURRENT FILING DATE: 2002-09-17  
; PRIOR APPLICATION NUMBER: US 60/323,469  
; PRIOR FILING DATE: 2001-09-17  
; NUMBER OF SEQ ID NOS: 412  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 205  
; LENGTH: 1211  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-29560A-205

Query Match 50.0%; Score 35; DB 1; Length 1211;  
Best Local Similarity 55.6%; Pred. No. 3.9e+02;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 3 HRIHESAS 11  
Db 444 HRFHWSRCS 452

RESULT 36  
US-60-487-610-1551  
; Sequence 1551, Application US/60487610  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: HUANG, Hongjin  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,  
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01469  
; CURRENT APPLICATION NUMBER: US/60/487,610  
; CURRENT FILING DATE: 2003-07-17  
; NUMBER OF SEQ ID NOS: 97101  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1551  
; LENGTH: 1211  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-487-610-1551

Query Match 50.0%; Score 35; DB 7; Length 1211;  
Best Local Similarity 55.6%; Pred. No. 3.9e+02;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 3 HRIHESAS 11  
Db 444 HRFHWSRCS 452

RESULT 37  
US-10-612-783-6253  
; Sequence 6253, Application US/10612783  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53373)A

CURRENT APPLICATION NUMBER: US/10/612,783  
CURRENT FILING DATE: 2003-07-02  
NUMBER OF SEQ ID NOS: 7098  
SEQ ID NO 6253  
LENGTH: 210  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(210)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4577\_78571C.1.pcp  
US-10-612-783-6253

Query Match 48.6%; Score 34; DB 6; Length 210;  
Best Local Similarity 62.5%; Pred. No. 98;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 HRIHWESA 10  
DB 175 HRFHWACH 182

RESULT 38  
US-09-674-546A-778  
Sequence 778, Application US/09674546A  
GENERAL INFORMATION:  
APPLICANT: Institute for Genomic Research  
APPLICANT: Chiron Corporation  
TITLE OF INVENTION: Neisseria Meningitidis Antigens and Compositions  
FILE REFERENCE: CHIR-0334  
CURRENT APPLICATION NUMBER: US/09/674,546A  
CURRENT FILING DATE: 1999-04-30  
NUMBER OF SEQ ID NOS: 3264  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 778  
LENGTH: 213  
TYPE: PRT  
ORGANISM: Neisseria meningitidis  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (23)..(23)  
OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (35)..(35)  
OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
US-09-674-546A-778

Query Match 48.6%; Score 34; DB 5; Length 213;  
Best Local Similarity 38.5%; Pred. No. 1e+02;  
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 ITHRIHWESASL 13  
DB 31 LTRXHWKMPSTAL 43

RESULT 39  
US-10-637-011-21  
Sequence 21, Application US/10637011  
GENERAL INFORMATION:  
APPLICANT: JESTIN, Andre  
APPLICANT: ALBINA, Emmanuel  
APPLICANT: Le CANN, Pierre  
APPLICANT: BLANCHARD, Phillipe  
APPLICANT: HUTET, Evelyne  
APPLICANT: ARNAUD, Claire  
APPLICANT: TRUONG, Catherine  
APPLICANT: MAHE, Dominique  
APPLICANT: CARIOLET, Roland  
APPLICANT: MADEC, Francois

TITLE OF INVENTION: CIRCOVIRUS SEQUENCES ASSOCIATED WITH PIGLET WEIGHT LOSS  
TITLE OF INVENTION: DISEASE (PMD)  
FILE REFERENCE: 065691/0176  
CURRENT APPLICATION NUMBER: US/10/637,011  
CURRENT FILING DATE: 2003-08-08  
PRIOR APPLICATION NUMBER: US/09/514,245B  
PRIOR FILING DATE: 2000-02-28  
PRIOR APPLICATION NUMBER: FR 97/15396  
PRIOR FILING DATE: 1997-12-05  
NUMBER OF SEQ ID NOS: 170  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 21  
LENGTH: 566  
TYPE: PRT  
ORGANISM: Type B PMD circovirus  
US-10-637-011-21

Query Match 48.6%; Score 34; DB 6; Length 566;  
Best Local Similarity 40.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 HRIHWESA 12  
DB 510 HOLHWREKEL 519

RESULT 40  
US-10-613-520-797  
Sequence 797, Application US/10613520  
GENERAL INFORMATION:  
APPLICANT: Monsanto Technology, LLC  
APPLICANT: Laurie, Cathy C  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants  
FILE REFERENCE: 38-77(52900)E  
CURRENT APPLICATION NUMBER: US/10/613,520  
CURRENT FILING DATE: 2003-07-02  
PRIOR APPLICATION NUMBER: US 60/365,301  
PRIOR FILING DATE: 2002-03-15  
PRIOR APPLICATION NUMBER: US 10/389,566  
PRIOR FILING DATE: 2003-03-14  
PRIOR APPLICATION NUMBER: US 60/391,786  
PRIOR FILING DATE: 2002-06-25  
PRIOR APPLICATION NUMBER: US 60/392,018  
PRIOR FILING DATE: 2002-06-26  
NUMBER OF SEQ ID NOS: 2578  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 797  
LENGTH: 694  
TYPE: PRT  
ORGANISM: Oryza sativa  
US-10-613-520-797

Query Match 48.6%; Score 34; DB 6; Length 694;  
Best Local Similarity 50.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 RIHWESASL 13  
DB 44 RIHWERGLV 53

Search completed: August 28, 2003, 14:21:50  
Job time : 24 secs

**THIS PAGE BLANK (USPTO)**

PT Biopolymer marker useful for evidencing, categorizing or regulating at

PT Infarction or intracerebral hemorrhage -  
 XX  
 PS Claim 1, Page 30; 30pp; English.  
 CC The present sequence is that of a biopolymer marker that is useful  
 CC in indicating at least one particular disease state. The marker  
 CC is a C3f fragment from the complement system having a mol. wt. of  
 CC about 1690 Da. The marker is indicative of an individual suffering  
 CC from myocardial infarction, intracerebral hemorrhage or congestive  
 CC heart failure. The invention involves the use of a combination of  
 CC preparatory steps in conjunction with mass spectroscopy and  
 CC time-of-flight detection procedures to maximise the diversity of  
 CC biopolymers which are verifiable within a particular sample. The  
 CC cohort of biopolymers verified within such a sample is then viewed  
 CC with reference to their ability to evidence at least one particular  
 CC disease state, thereby enabling a diagnostician to gain the ability  
 CC to characterise either the presence or absence of a disease state  
 CC relative to recognition of the presence and/or absence of the  
 CC biopolymer. Diagnostic assay kits including the present biopolymer  
 CC marker are claimed, and a process for regulating a disease state  
 CC by controlling the presence or absence of the biopolymer marker is  
 CC also claimed.  
 CC  
 SQ Sequence 14 AA;  
 Query Match 100.0%; Score 70; DB 24; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 3,4e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 ITHRIHWSASLL 13  
 |||||  
 Db 2 ITHRIHWSASLL 14  
 |||||  
 RESULT 2  
 ABG75803  
 ID ABG75803 standard; Peptide; 15 AA.  
 XX  
 AC ABG75803;  
 XX  
 DT 07-MAY-2003 (first entry)  
 XX  
 DE Congestive heart failure C3f fragment marker biopolymer.  
 XX  
 KW Biopolymer; disease state; congestive heart failure; asymptomatic;  
 KW diabetes; kidney failure; heart failure; Syndrome X; heart attack;  
 KW stroke; C3f; complement system.  
 XX  
 OS Unidentified.  
 OS  
 PN US2002160419-A1.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 30-APR-2001; 2001US-0845739.  
 XX  
 PR 30-APR-2001; 2001US-0845739.  
 XX  
 PA (JACK/) JACKOWSKI G.  
 PA (THAT/) THATCHER B.  
 PA (MARS/) MARSHALL J.  
 PA (YANT/) YANTHA J.  
 PA (VREE/) VREES T.  
 XX  
 PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;  
 XX  
 DR WPI; 2003-246643/25.  
 XX  
 PT A novel biopolymer marker useful in indicating at least one particular  
 PT disease state particularly congestive heart failure -  
 XX  
 PS Claim 1, Page 7; 10pp; English.

CC The invention discloses a biopolymer marker which is useful in indicating  
 CC at least one particular disease state, particularly congestive heart  
 CC failure. The polymer marker is also useful for characterising disease  
 CC states. The polymer marker allows physicians to identify asymptomatic  
 CC patients before they suffer an irreversible event such as diabetes,  
 CC kidney failure and heart failure, and enable effective disease management  
 CC and preventative medicine. Additionally, the specific diagnostic tests  
 CC which evolve using the biopolymer provide a tool for rapidly and  
 CC accurately diagnosing acute Syndrome X such as heart attack and stroke,  
 CC and facilitate treatment. The sequence presented is the congestive heart  
 CC failure specific C3f fragment from the complement system, which has a  
 CC molecular weight of about 1793 Daltons, marker biopolymer.  
 CC  
 SQ Sequence 15 AA;  
 Query Match 100.0%; Score 70; DB 24; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 3,7e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 ITHRIHWSASLL 13  
 |||||  
 Db 3 ITHRIHWSASLL 15  
 |||||  
 RESULT 3  
 ABU0617  
 ID ABU0617 standard; peptide; 17 AA.  
 XX  
 AC ABU0617;  
 XX  
 DT 23-MAY-2003 (first entry)  
 XX  
 DE Disease specific biopolymer marker #1.  
 XX  
 KW Biopolymer marker; type II diabetes; immunoassay.  
 XX  
 OS Homo sapiens.  
 OS  
 PN US2002160532-A1.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 30-APR-2001; 2001US-0846346.  
 XX  
 PR 30-APR-2001; 2001US-0846346.  
 XX  
 PA (JACK/) JACKOWSKI G.  
 PA (THAT/) THATCHER B.  
 PA (MARS/) MARSHALL J.  
 PA (YANT/) YANTHA J.  
 PA (VREE/) VREES T.  
 XX  
 PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;  
 XX  
 DR WPI; 2003-328370/31.  
 XX  
 PT Biopolymer marker useful in indicating disease state, in particular  
 PT type II diabetes and as antigens in immunoassays for detecting  
 PT individuals suffering from disease known to be evidenced by marker  
 PT sequence -  
 XX  
 PS Claim 1, Page 7; 10pp; English.  
 XX  
 CC The invention describes a biopolymer marker (I) useful in indicating at  
 CC least one particular disease state. (I) is useful for indicating a  
 CC disease state, in particular type II diabetes. The marker sequences are  
 CC useful as antigens in immunoassays for the detection of those individuals  
 CC suffering from the disease known to be evidenced by the marker sequence.  
 CC (II) provides an efficient diagnostic tool for rapidly and accurately  
 CC diagnosing disease states such as type II diabetes. This is the  
 CC amino acid sequence of a biopolymer marker.  
 CC  
 SQ Sequence 17 AA;

Query Match 100.0%; Score 70; DB 24; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ITHRIHWESASLL 13  
 |||||  
 DB 4 ITHRIHWESASLL 16

## RESULT 4

AB82769  
 ID AB82769 standard; peptide; 17 AA.

XX AC AB82769;

XX DT 18-MAR-2003 (first entry)

XX DE Congestive heart failure indicative biopolymer marker.

XX KW Biopolymer; marker; C3f; complement system; congestive heart failure;  
 human.

XX OS Homo sapiens.

XX PN WO200288717-A2.

XX PD 07-NOV-2002.

XX PP 25-APR-2002; 2002WO-CA00578.

XX PR 30-APR-2001; 2001US-0845736.

XX PA (SYNX-) SYN.X PHARMA INC.

XX PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;

XX DR WPI; 2003-120486/11.

XX PT Use of biopolymer marker for evidencing, categorizing or regulating at  
 least one disease state, e.g. congestive heart failure

XX PS Claim 1; Fig 1; 27pp; English.

XX CC The present sequence represents a biopolymer marker of the invention and  
 is a disease specific marker. The marker is characterised as a C3f  
 fragment from the complement system having a molecular weight of about  
 2056 daltons. The biopolymer marker identified is useful for evidencing,  
 categorizing or regulating at least one disease state, preferably  
 congestive heart failure.

XX SQ Sequence 17 AA;

Query Match 100.0%; Score 70; DB 24; Length 17;

Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ITHRIHWESASLL 13  
 |||||  
 DB 4 ITHRIHWESASLL 16

## RESULT 5

ABU08618  
 ID ABU08618 standard; peptide; 18 AA.

XX AC ABU08618;

XX DT 23-MAY-2003 (first entry)

XX DE Disease specific biopolymer marker #2.

XX KW Biopolymer marker; type II diabetes; immunoassay.

XX OS Homo sapiens.  
 XX XX US2002160532-A1.

XX PD 31-OCT-2002.

XX PF 30-APR-2001; 2001US-0846346.

XX PR 30-APR-2001; 2001US-0846346.

XX PA (JACK/) JACKOWSKI G.

XX PA (THAT/) THATCHER B.

XX PA (MARS/) MARSHALL J.

XX PA (YANT/) YANTHA J.

XX PA (VREE/) VREES T.

XX PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;

XX DR WPI; 2003-328370/31.

XX PS Disclosure; Fig 2; 10pp; English.

XX CC The invention describes a biopolymer marker (1) useful in indicating at  
 least one particular disease state. (1) is useful for indicating a  
 disease state, in particular type II diabetes. The marker sequences are  
 useful as antigens in immunoassays for the detection of those individuals  
 suffering from the disease known to be evidenced by the marker sequence.  
 CC (1) provides an efficient diagnostic tool for rapidly and accurately  
 diagnosing disease states such as type II diabetes. This is the  
 amino acid sequence of a biopolymer marker.

XX SQ Sequence 18 AA;

Query Match 100.0%; Score 70; DB 24; Length 18;

Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ITHRIHWESASLL 13  
 |||||  
 DB 5 ITHRIHWESASLL 17

## RESULT 6

ABG25976  
 ID ABG25976 standard; Protein; 1540 AA.

XX AC ABG25976;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #25967.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;  
XX  
XX MPI: 2001-639362/73.  
DR N-PSDB; AAS90163.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostic, forensic, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
XX Claim 20; SEQ ID No 56335; 103bp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABC00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 1540 AA;  
Query Match 100.0%; Score 70; DB 22; Length 1540;  
Best Local Similarity 100.0%; Pred. No. 0.0064;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ITHRIHWSASLL 13  
Db 1307 ITHRIHWSASLL 1319  
RESULT 7  
AAW34623 standard; Protein; 1592 AA.  
XX  
XX AAW34623;  
AC  
XX  
DT 09-APR-1998 (first entry)  
XX  
XX Human C3 protein mutant FT-1.  
DE  
XX  
XX Human; C3 protein; convertase; complement pathway protein; infection;  
KW down-regulation resistant C3 convertase; xenograft rejection; therapy;  
KW complement-mediated disease; autoimmune disease; leukaemia cell; tumour;  
KW complement-mediated response; MHC-mismatched lymphocyte; mutein.  
XX  
XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1591 /note= "R1591T mutation"  
FT Misc-difference 1592 /note= "R1592N mutation"  
FT Misc-difference 1593 /note= "A1593Stop mutation"  
XX  
XX W09732981-A1.  
XX  
XX

PD 12-SEP-1997.  
XX  
XX 04-MAR-1997; 97WO-GB00603.  
XX  
XX 19-NOV-1996; 96GB-0024028.  
XX  
XX 07-MAR-1996; 96GB-0004865.  
PR 07-JUN-1996; 96GB-0011896.  
PR 08-JUL-1996; 96GB-0014293.  
XX  
XX (IMUT-) IMUTRAN LTD.  
XX  
XX Farries TC, Harrison RA;  
PI  
XX  
XX MPI, 1997-457534/42.  
XX  
XX Modified complement pathway protein that forms C3 convertase  
PT resistant to down-regulation - used to exhaust the complement  
PT pathway by super-activation, especially for preventing graft  
PT rejection, etc.  
XX  
XX Example 17; Page -; 123pp; English.  
XX  
XX This sequence represents a mutated human C3 protein of the invention  
CC (see AAW34606 for wild type protein). This protein is a protein of the  
CC invention, and is a modified native complement pathway protein (A) that  
CC forms a down-regulation resistant C3 convertase. (A), their variants,  
CC fragments and conjugates are used to deplete levels of complement  
CC pathway proteins (by superactivation until one or more components are  
CC exhausted), specifically to prevent rejection of foreign material  
CC (particularly a xenograft) but also to prevent complement-mediated  
CC diseases resulting from (surgical) injury or antibody-antigen interaction  
CC in autoimmune disease, also to localise and/or amplify endogenous  
CC complement protein conversion and deposition at a specific site (e.g. a  
CC virus, infected cell or tumour, to increase sensitivity to  
CC complement-mediated responses; a particular application is eliminating  
CC any cancer cells left after surgical removal of a tumour). Also  
CC contemplated is ex vivo treatment, especially by passing blood through a  
CC matrix containing (A) (this may remove additional anaphylactic peptides  
CC and other inflammatory mediators) or killing of leukaemia cells or  
CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
CC inactivated), causing inactivation of the alternative pathway by  
CC consumption of factor B.  
XX  
XX  
SQ Sequence 1592 AA;  
Query Match 100.0%; Score 70; DB 18; Length 1592;  
Best Local Similarity 100.0%; Pred. No. 0.0066;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ITHRIHWSASLL 13  
Db 1307 ITHRIHWSASLL 1319  
RESULT 8  
AAW34624 standard; Protein; 1635 AA.  
XX  
XX AAW34624;  
AC  
XX  
DT 09-APR-1998 (first entry)  
XX  
XX Human C3 protein mutant FT-2.  
DE  
XX  
XX Human; C3 protein; convertase; complement pathway protein; infection;  
KW down-regulation resistant C3 convertase; xenograft rejection; therapy;  
KW complement-mediated disease; autoimmune disease; leukaemia cell; tumour;  
KW complement-mediated response; MHC-mismatched lymphocyte; mutein.  
XX  
XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
XX  
XX



KM		complement-mediated response; MHC-mismatched lymphocyte; mutein.
XX		
OQ	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference	1638..1645
PT	/note= "wild type residues QDENQK mutated to SS"	
PN	WO9732981-A1.	
PD	12-SEP-1997.	
XX		
PF	04-MAR-1997;	97WO-GB00603.
PR	19-NOV-1996;	96GB-0024028.
XX		
PR	07-MAR-1996;	96GB-0004865.
PR	07-JUN-1996;	96GB-0011896.
PR	08-JUL-1996;	96GB-0014293.
XX		
PA	(IMUT-) IMUTRAN LTD.	
PI	Farries TC, Harrison RA;	
XX		
DR	WPI; 1997-457534/42.	
XX		
PT	Modified complement pathway protein that forms C3 convertase	
PT	resistant to down-regulation - used to exhaust the complement	
PT	pathway by super-activation, especially for preventing graft	
PT	rejection, etc.	
PS		
PX	Example 17; Page -: 123pp; English.	
XX		
CC	This sequence represents a mutated human C3 protein of the invention	
CC	(see AAMW34606 for wild type protein). This protein is a protein of the	
CC	invention, and is a modified native complement pathway protein (A) that	
CC	forms a down-regulation resistant C3 convertase. (A), their variants,	
CC	fragments and conjugates are used to deplete levels of complement	
CC	pathway proteins (by superactivation until one or more components are	
CC	exhausted), specifically to prevent rejection of foreign material	
CC	(particularly a xenograft) but also to prevent complement-mediated	
CC	diseases resulting from (surgical) injury or antibody-antigen interaction	
CC	in autoimmune disease, also to localise and/or amplify endogenous	
CC	complement protein conversion and deposition at a specific site (e.g. a	
CC	virus, infected cell or tumour, to increase sensitivity to	
CC	complement-mediated responses; a particular application is eliminating	
CC	any cancer cells left after surgical removal of a tumour). Also	
CC	contemplated is ex vivo treatment, especially by passing blood through a	
CC	matrix containing (A) (this may remove additional anaphylactic peptides	
CC	and other inflammatory mediators) or killing of leukaemia cells or	
CC	MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not	
CC	inhibited by factor I, it can bind repeatedly to factor B (which is then	
CC	inactivated), causing inactivation of the alternative pathway by	
CC	consumption of factor B.	
SQ	Sequence	1657 AA;
Query Match	100.0%; Score 70; DB 18; Length 1657;	
Best Local Similarity	100.0%; Pred. No. 0.0069;	
Matches	13; Conservative	0; Mismatches
OY	1 ITHRIHWESASLL 13	Indels
DB	1307 ITHRIHWESASLL 1319	Gaps
RESULT 10		
AAMW34625		
ID	AAMW34625 standard; Protein; 1661 AA.	
AC	AAMW34625;	
DT	09-APR-1998 (first entry)	

```

DE Human C3 protein mutant FT-3.
XX
XX
XX Human; C3 protein; convertase; complement pathway protein; infection;
XX down-regulation resistant C3 convertase; xenograft rejection; therapy;
XX complement-mediated disease; autoimmune disease; leukemia cell; tumour;
XX complement-mediated response; MHC-mismatched lymphocyte; mutein.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX MISC-difference 1607..1614
XX /note="Wild type residues LSSDFWGE mutated to KEALQI"
XX
XX WO9732981-A1.
XX
XX 12-SEP-1997.
XX
XX 04-MAR-1997; 97WO-GB00603.
XX
XX 19-NOV-1996; 96GB-0024028.
XX 07-MAR-1996; 96GB-0004865.
XX 07-JUN-1996; 96GB-0011896.
XX 08-JUL-1996; 96GB-0014293.
XX
XX (IMUT-) IMUTRAN LTD.
XX
XX Farries TC, Harrison RA;
XX
XX WPI; 1997-457534/42.
XX
XX Modified complement pathway protein that forms C3 convertase
XX resistant to down-regulation - used to exhaust the complement
XX pathway by super-activation, especially for preventing graft
XX rejection, etc.
XX
XX Example 17, Page -: 123pp; English.
XX
XX This sequence represents a mutated human C3 protein of the invention
XX (see AAW34606 for wild type protein). This protein is a protein of the
XX invention, and is a modified native complement pathway protein (A) that
XX forms a down-regulation resistant C3 convertase. (A), their variants,
XX fragments and conjugates are used to deplete levels of complements
XX pathway proteins (by superactivation until one or more components are
XX exhausted), specifically to prevent rejection of foreign material
XX (particularly a xenograft) but also to prevent complement-mediated
XX diseases resulting from (surgical) injury or antibody-antigen interaction
XX in autoimmune disease, also to localise and/or amplify endogenous
XX complement protein conversion and deposition at a specific site (e.g. a
XX virus, infected cell or tumour, to increase sensitivity to
XX complement-mediated responses; a particular application is eliminating
XX any cancer cells left after surgical removal of a tumour). Also
XX contemplated is ex vivo treatment, especially by passing blood through a
XX matrix containing (A) (this may remove additional anaphylactic peptides
XX and other inflammatory mediators) or killing of leukaemia cells or
XX MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not
XX inhibited by factor I, it can bind repeatedly to factor B (which is then
XX inactivated), causing inactivation of the alternative pathway by
XX consumption of factor B.
XX
XX
XX Sequence 1661 AA;
XX
XX Query Match 100.0%; Score 70; DB 18; Length 1661;
XX Best Local Similarity 100.0%; Pred. NO. 0.0069;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ITHRHMESASLL 13
XX |||||
XX Db 1307 ITHRHMESASLL 1319
XX
XX RESULT 11
XX AAR94028
XX ID AAR94028 standard; Protein; 1663 AA.

```

```

XX
XX AAR94028;
XX
XX 21-MAY-1996 (first entry)
XX
XX Human C3 precursor.
XX
XX C3 protein; convertase; Factor I; Factor H; complement.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX FT Peptide 1..22
XX FT /label= Sig_peptide
XX FT 23..667
XX FT /note="C3 beta chain"
XX FT 668..671
XX FT /note="amino acids 668-671 are removed when the
XX FT precursor is cleaved into the alpha and
XX FT beta chains"
XX FT Protein 672..1663
XX FT /note="C3 alpha chain"
XX
XX WO9607738-A2.
XX
XX 14-MAR-1996.
XX
XX 08-SEP-1995; 95WO-GB02121.
XX
XX 04-MAY-1995; 95GB-0009102.
XX 08-SEP-1994; 94GB-0018147.
XX
XX (IMUT-) IMUTRAN LTD.
XX
XX Farries TC, Harrison RA;
XX
XX WPI; 1996-171613/17.
XX N-PsDB; AAT17738.
XX
XX Mutant complement pathway protein forming stable C3 convertase -
XX for generalised complement depletion or localised complement
XX activation
XX
XX Disclosure; Fig 1; 81pp; English.
XX
XX Human C3 protein (AAR94028) was produced by expression of a cDNA
XX sequence (AAT17738) isolated from a human liver cDNA library.
XX C3 is a complement pathway protein that is susceptible to cleavage
XX by factor I and is also susceptible to the inhibitory action
XX of factor H. Mutants of C3 (AAR94029 and AAR94030) have been
XX produced by site-directed mutagenesis. These mutants can be
XX used to super-activate the complement system, or to induce
XX localised super-activation at a specific target to increase
XX the target's sensitivity to complement-mediated destruction.
XX
XX
XX Sequence 1663 AA;
XX
XX Query Match 100.0%; Score 70; DB 17; Length 1663;
XX Best Local Similarity 100.0%; Pred. NO. 0.0069;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ITHRHMESASLL 13
XX |||||
XX Db 1307 ITHRHMESASLL 1319
XX
XX RESULT 12
XX AAR94029
XX ID AAR94029 standard; Protein; 1663 AA.
XX
XX AAR94029;
XX
XX 21-MAY-1996 (first entry)
XX

```

XX		Human modified C3 (R1303X).	
DE			
XX		C3 protein; convertase; Factor I; Factor H; complement;	
KW		tumour; infection; therapy.	
XV			
XX		Synthetic.	
OS			
XX			
FH	Key	Location/Qualifiers	
FT	Peptide	1..22	
FT		/label= Sig_peptide	
FT	Protein	23..667	
FT		/note= "C3 beta chain"	
FT	Peptide	668..671	
FT		/note= "amino acids 668-671 are removed when the precursor is cleaved into the alpha and beta chains"	
FT	Protein	672..1663	
FT		/note= "C3 alpha chain"	
FT	Misc-difference	1303	
FT		/label= Glu, Gly, Gln	
XX			
PN	MO9607738-A2.		
XX			
PD	14-MAR-1996.		
XX			
PP	08-SEP-1995;	95WO-GB02121.	
XX			
PR	04-MAY-1995;	95GB-0009102.	
PR	08-SEP-1994;	94GB-0018147.	
XX			
PA	(IMUT-) IMOTRAN LTD.		
XX			
P1	Farries TC, Harrison RA;		
XX			
DR	WPI; 1996-171613/17.		
XX			
PT	Mutant complement pathway protein forming stable C3 convertase -		
FT	for generalised complement depletion or localised complement		
PT	activation		
XX			
PS	Claim 6; Fig 1; 81pp; English.		
XX			
CC	A modified human C3 protein (AAR94029) differs from the wild-type		
CC	(AAR94028) by substitution of Arg-1303 by glutamic acid, glycine		
CC	or glutamine. It is obtained by site-directed mutagenesis of		
CC	C3-encoding cDNA (AA1717738). The modification results in improved		
CC	resistance to cleavage by Factor I in comparison to wild-type C3.		
CC	This allows the modified C3 to be used therapeutically to		
CC	super-activate the complement system or the increase a target's		
CC	(e.g. tumour, pathogen or virus-infected cell) sensitivity to		
CC	complement-mediated destruction.		
XX			
SO	Sequence 1663 AA;		
QY			
	Query Match 100.0%; Score 70; DB 17; Length 1663;		
	Best Local Similarity 100.0%; Pred. No. 0.0069;		
	Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
	1 ITHRIHWESASILL 13		
DB	1307 ITHRIHWESASILL 1319		
	RESULT 13		
ID	AAR94030		
XX	AAR94030 standard; Protein; 1663 AA.		
AC	AAR94030;		
XX			
DT	21-MAY-1996 (first entry)		
XX			
DE	Human modified C3 (D752G, E753S, D754G).		

XX	C3 protein; convertase; Factor I; Factor H; complement; tumour;
KM	Infection; therapy.
KM	Synthetic.
OS	
XX	
FH	Key
FT	Location/Qualifiers
FT	1..22
FT	/label= Sig_peptide
FT	23..667
FT	/note= "C3 beta chain"
FT	668..671
FT	/note= "amino acids 668-671 are removed when the precursor is cleaved into the alpha and beta chains"
FT	
FT	Protein
FT	672..1663
FT	/note= "C3 alpha chain"
PN	
PD	W09607738-A2.
PD	14-MAR-1996.
PF	
PF	08-SEP-1995; 95WO-GB02121.
PR	
PR	04-MAY-1995; 95GB-0009102.
PR	08-SEP-1994; 94GB-0018147.
PA	(IMOT-) IMUTRAN LTD.
XX	
P1	Farries TC; Harrison RA;
DR	WPI; 1996-171613/17.
XX	
PT	Mutant complement pathway protein forming stable C3 convertase -
PT	for generalised complement depletion or localised complement
PT	activation
XX	
PS	Claim 11; Fig 1; 81dp; English.
XX	
CC	A modified human C3 protein (AAR94030) differs from the wild-type
CC	(AAR94028) by substitution of Asp-Glu-Asp at positions 752-754 by
CC	Gly-Ser-Gly. It is obtained by site-directed mutagenesis of
CC	C3-encoding cDNA (AA117738). The modification reduces the
CC	interaction of C3b/C3i with Factor H in comparison to wild-type
CC	C3. This allows the modified C3 to be used therapeutically to
CC	super-active the complement system or the increase a target's
CC	(e.g. tumour, pathogen or virus-infected cell) sensitivity to
CC	complement-mediated destruction.
XX	
SQ	Sequence 1663 AA;
Query Match	100.0%; Score 70; DB 17; Length 1663;
Best Local Similarity	100.0%; Pred. NO. 0.0069;
Matches 13; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 ITHRIHWESASLT 13             
Db	1307 ITHRIHWESASLT 1319
RESULT 14	
ID AAW34619	AAW34619 standard; Protein; 1663 AA.
XX AAW34619;	
AC	
DT 09-APR-1998	(first entry)
XX	
DE Human C3 protein mutant DV-9.	
KM	Human, C3 protein; convertase; complement pathway protein; infection;
KM	down-regulation resistant C3 convertase; xenograft rejection; therapy;
KM	complement-mediated disease; autoimmune disease; leukaemia cell; tumour;

KW complement-mediated response; MHC-mismatched lymphocyte; mutein.  
 XX Homo sapiens.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1216  
 FT /note= "D1216G mutation"  
 FT Misc-difference 1217  
 FT /note= "K1217E mutation"  
 FT Misc-difference 1218  
 FT /note= "N1218D mutation"  
 FT Misc-difference 1219  
 FT /note= "R1219H mutation"  
 XX  
 XX MO9732981-A1.  
 XX  
 XX  
 PD 12-SEP-1997.  
 XX  
 PF 04-MAR-1997; 97WO-GB00603.  
 XX  
 XX 19-NOV-1996; 96GB-0024028.  
 PR 07-MAR-1996; 96GB-0004865.  
 PR 07-JUN-1996; 96GB-0011896.  
 PR 08-JUL-1996; 96GB-0014293.  
 XX  
 XX (IMUT-) IMUTRAN LTD.  
 XX  
 XX Farries TC, Harrison RA;  
 XX  
 DR WPI, 1997-457534/42.  
 XX  
 PT Modified complement pathway protein that forms C3 convertase  
 PT resistant to down-regulation - used to exhaust the complement  
 PT pathway by super-activation, especially for preventing graft  
 PT rejection, etc.  
 XX  
 XX Example 14; Page -: 123p; English.  
 XX  
 XX This sequence represents a mutated human C3 protein of the invention  
 CC (see AAM34620 for wild type protein). This protein is a protein of the  
 CC invention, and is a modified native complement pathway protein (A) that  
 CC forms a down-regulation resistant C3 convertase. (A), their variants,  
 CC fragments and conjugates are used to deplete levels of complement  
 CC pathway proteins (by superactivation until one or more components are  
 CC exhausted), specifically to prevent rejection of foreign material  
 CC (particularly a xenograft) but also to prevent complement-mediated  
 CC diseases resulting from (surgical) injury or antibody-antigen interaction  
 CC in autoimmune disease, also to localise and/or amplify endogenous  
 CC complement protein conversion and deposition at a specific site (e.g. a  
 CC virus, infected cell or tumour, to increase sensitivity to  
 CC complement-mediated responses; a particular application is eliminating  
 CC any cancer cells left after surgical removal of a tumour). Also  
 CC contemplated is ex vivo treatment, especially by passing blood through a  
 CC matrix containing (A) (this may remove additional anaphylactic peptides  
 CC and other inflammatory mediators) or killing of leukaemia cells or  
 CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
 CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
 CC inactivated), causing inactivation of the alternative pathway by  
 CC consumption of factor B.  
 CC  
 XX  
 XX Sequence 1663 AA;  
 SQ  
 Query Match 100.0%; Score 70; DB 18; Length 1663;  
 Best Local Similarity 100.0%; Pred. No. 0.0069;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ITHRIHMSASLL 13  
 |||||  
 Db 1307 ITHRIHMSASLL 1319  
 RESULT 15  
 AAM34620

ID AAM34620 standard; Protein, 1663 AA.  
 XX  
 XX AAM34620;  
 AC  
 XX  
 XX 09-APR-1998 (first entry)  
 DT  
 XX  
 XX Human C3 protein mutant CV-4.  
 DE  
 XX  
 XX Human; C3 protein; convertase; complement pathway protein; infection;  
 KW down-regulation resistant C3 convertase; xenograft rejection; therapy;  
 KW complement-mediated disease; autoimmune disease; leukaemia cell; tumour;  
 KW complement-mediated response; MHC-mismatched lymphocyte; mutein.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1260  
 FT /note= "R1260N mutation"  
 FT Misc-difference 1264  
 FT /note= "G1264E mutation"  
 FT  
 XX  
 XX MO9732981-A1.  
 XX  
 XX  
 PD 12-SEP-1997.  
 XX  
 PF 04-MAR-1997; 97WO-GB00603.  
 XX  
 XX 19-NOV-1996; 96GB-0024028.  
 PR 07-MAR-1996; 96GB-0004865.  
 PR 07-JUN-1996; 96GB-0011896.  
 PR 08-JUL-1996; 96GB-0014293.  
 XX  
 XX (IMUT-) IMUTRAN LTD.  
 XX  
 XX Farries TC, Harrison RA;  
 XX  
 DR WPI, 1997-457534/42.  
 XX  
 PT Modified complement pathway protein that forms C3 convertase  
 PT resistant to down-regulation - used to exhaust the complement  
 PT pathway by super-activation, especially for preventing graft  
 PT rejection, etc.  
 XX  
 XX Example 14; Page -: 123p; English.  
 XX  
 XX This sequence represents a mutated human C3 protein of the invention  
 CC (see AAM34620 for wild type protein). This protein is a protein of the  
 CC invention, and is a modified native complement pathway protein (A) that  
 CC forms a down-regulation resistant C3 convertase. (A), their variants,  
 CC fragments and conjugates are used to deplete levels of complement  
 CC pathway proteins (by superactivation until one or more components are  
 CC exhausted), specifically to prevent rejection of foreign material  
 CC (particularly a xenograft) but also to prevent complement-mediated  
 CC diseases resulting from (surgical) injury or antibody-antigen interaction  
 CC in autoimmune disease, also to localise and/or amplify endogenous  
 CC complement protein conversion and deposition at a specific site (e.g. a  
 CC virus, infected cell or tumour, to increase sensitivity to  
 CC complement-mediated responses; a particular application is eliminating  
 CC any cancer cells left after surgical removal of a tumour). Also  
 CC contemplated is ex vivo treatment, especially by passing blood through a  
 CC matrix containing (A) (this may remove additional anaphylactic peptides  
 CC and other inflammatory mediators) or killing of leukaemia cells or  
 CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
 CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
 CC inactivated), causing inactivation of the alternative pathway by  
 CC consumption of factor B.  
 CC  
 XX  
 XX Sequence 1663 AA;  
 SQ  
 Query Match 100.0%; Score 70; DB 18; Length 1663;  
 Best Local Similarity 100.0%; Pred. No. 0.0069;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ITHRIHWEASL 13  
 |||||  
 Db 1307 ITHRIHWEASL 1319

RESULT 16  
 AAW34621  
 ID AAW34621 standard; Protein; 1663 AA.

AC AAW34621;

DT 09-APR-1998 (first entry)

DE Human C3 protein mutant RY-1.

Human; C3 protein; convertase; complement pathway protein; infection;  
 down-regulation resistant C3 convertase; xenograft rejection; therapy;  
 complement-mediated disease; autoimmune disease; leukaemia cell; tumour;  
 complement-mediated response; MHC-mismatched lymphocyte; mutein.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 1427

FT Misc-difference /note= "R1427Q mutation"

FT Misc-difference 1431

FT Misc-difference /note= "K1431D mutation"

FT Misc-difference 1433

FT Misc-difference /note= "E1433Q mutation"

PN WO9732981-A1.

PD 12-SEP-1997.

XX 04-MAR-1997; 97WO-GB00603.

XX 19-NOV-1996; 96GB-0024028.

PR 07-MAR-1996; 96GB-0004865.

PR 07-JUN-1996; 96GB-0011896.

PR 08-JUL-1996; 96GB-0014293.

XX (IMUT-) IMUTRAN LTD.

PI Farries TC, Harrison RA;

XX WPI; 1997-457534/42.

Modified complement pathway protein that forms C3 convertase  
 resistant to down-regulation - used to exhaust the complement  
 pathway by super-activation, especially for preventing graft  
 rejection, etc.

Example 14; Page -; 123pp; English.

This sequence represents a mutated human C3 protein of the invention  
 (see AAW34606 for wild type protein). This protein is a protein of the  
 invention, and is a modified native complement pathway protein (A) that  
 forms a down-regulation resistant C3 convertase. (A), their variants,  
 fragments and conjugates are used to deplete levels of complement  
 pathway proteins (by superactivation until one or more components are  
 exhausted), specifically to prevent rejection of foreign material  
 (particularly a xenograft) but also to prevent complement-mediated  
 diseases resulting from (surgical) injury or antibody-antigen interaction  
 in autoimmune disease, also to localise and/or amplify endogenous  
 complement protein conversion and deposition at a specific site (e.g. a  
 virus, infected cell or tumour, to increase sensitivity to  
 complement-mediated responses; a particular application is eliminating  
 any cancer cells left after surgical removal of a tumour). Also  
 contemplated is ex vivo treatment, especially by passing blood through a  
 matrix containing (A) (this may remove additional anaphylactic peptides  
 and other inflammatory mediators) or killing of leukaemia cells or  
 MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
 inhibited by factor I, it can bind repeatedly to factor B (which is then

CC inactivated), causing inactivation of the alternative pathway by  
 CC consumption of factor B.

SO Sequence 1663 AA;

Query Match 100.0%; Score 70; DB 18; Length 1663;  
 Best Local Similarity 100.0%; Pred. No. 0.0069;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ITHRIHWEASL 13  
 |||||  
 Db 1307 ITHRIHWEASL 1319

RESULT 17

AAW34627  
 ID AAW34627 standard; Protein; 1663 AA.

AC AAW34627;

DT 09-APR-1998 (first entry)

DE Human C3 protein mutant FT-5.

Human; C3 protein; convertase; complement pathway protein; infection;  
 down-regulation resistant C3 convertase; xenograft rejection; therapy;  
 complement-mediated disease; autoimmune disease; leukaemia cell; tumour;  
 complement-mediated response; MHC-mismatched lymphocyte; mutein.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 1661

FT Misc-difference /note= "C1661S mutation"

PN WO9732981-A1.

PD 12-SEP-1997.

XX 04-MAR-1997; 97WO-GB00603.

XX 19-NOV-1996; 96GB-0024028.

PR 07-MAR-1996; 96GB-0004865.

PR 07-JUN-1996; 96GB-0011896.

PR 08-JUL-1996; 96GB-0014293.

XX (IMUT-) IMUTRAN LTD.

PI Farries TC, Harrison RA;

XX WPI; 1997-457534/42.

Modified complement pathway protein that forms C3 convertase  
 resistant to down-regulation - used to exhaust the complement  
 pathway by super-activation, especially for preventing graft  
 rejection, etc.

Example 17; Page -; 123pp; English.

This sequence represents a mutated human C3 protein of the invention  
 (see AAW34606 for wild type protein). This protein is a protein of the  
 invention, and is a modified native complement pathway protein (A) that  
 forms a down-regulation resistant C3 convertase. (A), their variants,  
 fragments and conjugates are used to deplete levels of complement  
 pathway proteins (by superactivation until one or more components are  
 exhausted), specifically to prevent rejection of foreign material  
 (particularly a xenograft) but also to prevent complement-mediated  
 diseases resulting from (surgical) injury or antibody-antigen interaction  
 in autoimmune disease, also to localise and/or amplify endogenous  
 complement protein conversion and deposition at a specific site (e.g. a  
 virus, infected cell or tumour, to increase sensitivity to  
 complement-mediated responses; a particular application is eliminating  
 any cancer cells left after surgical removal of a tumour). Also

CC contemplated is ex vivo treatment, especially by passing blood through a  
CC matrix containing (A) (this may remove additional anaphylactic peptides  
CC and other inflammatory mediators) or killing of leukaemia cells or  
CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
CC inactivated), causing inactivation of the alternative pathway by  
CC consumption of factor B.  
XX  
SQ Sequence 1663 AA;  
Query Match 100.0%; Score 70; DB 18; Length 1663;  
Best Local Similarity 100.0%; Pred. No. 0.0069;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ITHRHWSASLL 13  
Db 1307 ITHRHWSASLL 1319  
RESULT 18  
AAM34628  
ID AAM34628 standard; Protein; 1663 AA.  
XX  
AC AAM34628;  
XX  
DT 09-APR-1998 (first entry)  
XX  
DE Human C3 protein mutant FR-2.  
XX  
Human; C3 protein; convertase; complement pathway protein; infection;  
KM down-regulation resistant C3 convertase; xenograft rejection; therapy;  
KW complement-mediated disease; autoimmune disease; leukaemia cell; tumour;  
KM complement-mediated response; MHC-mismatched lymphocyte; mutein.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1633 /note= "E1633R mutation"  
FT Misc-difference 1634 /note= "E1634D mutation"  
FT Misc-difference 1635 /note= "D1635T mutation"  
FT Misc-difference 1636 /note= "D1635T mutation"  
FT Misc-difference 1636 /note= "E1636T mutation"  
XX  
FN WO9732981-A1.  
XX  
PD 12-SEP-1997.  
XX  
PF 04-MAR-1997; 97WO-GB00603.  
XX  
PR 19-NOV-1996; 96GB-0024028.  
PR 07-MAR-1996; 96GB-0004865.  
PR 07-JUN-1996; 96GB-0011896.  
PR 08-JUL-1996; 96GB-0014293.  
XX  
PA (IMUT-) IMUTRAN LTD.  
XX  
PI Farries TC, Harrison RA;  
XX  
DR WPI, 1997-457534/42.  
XX  
PT Modified complement pathway protein that forms C3 convertase  
PT resistant to down-regulation - used to exhaust the complement  
PT pathway by super-activation, especially for preventing graft  
PT rejection, etc.  
XX  
PS Example 17; Page -; 123pp; English.  
XX  
CC This sequence represents a mutated human C3 protein of the invention  
CC (see AAM34606 for wild type protein). This protein is a protein of the  
CC invention, and is a modified native complement pathway protein (A) that

CC forms a down-regulation resistant C3 convertase. (A), their variants,  
CC fragments and conjugates are used to deplete levels of complement  
CC pathway proteins (by superactivation until one or more components are  
CC exhausted), specifically to prevent rejection of foreign material  
CC (particularly a xenograft) but also to prevent complement-mediated  
CC diseases resulting from (surgical) injury or antibody-antigen interaction  
CC in autoimmune disease, also to localise and/or amplify endogenous  
CC complement protein conversion and deposition at a specific site (e.g. a  
CC virus, infected cell or tumour, to increase sensitivity to  
CC complement-mediated responses; a particular application is eliminating  
CC any cancer cells left after surgical removal of a tumour). Also  
CC contemplated is ex vivo treatment, especially by passing blood through a  
CC matrix containing (A) (this may remove additional anaphylactic peptides  
CC and other inflammatory mediators) or killing of leukaemia cells or  
CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
CC inactivated), causing inactivation of the alternative pathway by  
CC consumption of factor B.  
XX  
SQ Sequence 1663 AA;  
Query Match 100.0%; Score 70; DB 18; Length 1663;  
Best Local Similarity 100.0%; Pred. No. 0.0069;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ITHRHWSASLL 13  
Db 1307 ITHRHWSASLL 1319  
RESULT 19  
AAM34630  
ID AAM34630 standard; Protein; 1663 AA.  
XX  
AC AAM34630;  
XX  
DT 09-APR-1998 (first entry)  
XX  
DE Human C3 protein mutant FR-3.  
XX  
Human; C3 protein; convertase; complement pathway protein; infection;  
KM down-regulation resistant C3 convertase; xenograft rejection; therapy;  
KW complement-mediated disease; autoimmune disease; leukaemia cell; tumour;  
KM complement-mediated response; MHC-mismatched lymphocyte; mutein.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1638..1645 /note= "wild type residues QDENQKQ mutated to RSTRQRAA"  
FT FT  
XX  
FN WO9732981-A1.  
XX  
PD 12-SEP-1997.  
XX  
PF 04-MAR-1997; 97WO-GB00603.  
XX  
PR 19-NOV-1996; 96GB-0024028.  
PR 07-MAR-1996; 96GB-0004865.  
PR 07-JUN-1996; 96GB-0011896.  
PR 08-JUL-1996; 96GB-0014293.  
XX  
PA (IMUT-) IMUTRAN LTD.  
XX  
PI Farries TC, Harrison RA;  
XX  
DR WPI, 1997-457534/42.  
XX  
PT Modified complement pathway protein that forms C3 convertase  
PT resistant to down-regulation - used to exhaust the complement  
PT pathway by super-activation, especially for preventing graft  
PT rejection, etc.  
XX

PS Example 17; Page -; 123pp; English.

CC This sequence represents a mutated human C3 protein of the invention  
 CC (see AAMW3606 for wild type protein). This protein is a protein of the  
 CC invention, and is a modified native complement pathway protein (A) that  
 CC forms a down-regulation resistant C3 convertase. (A), their variants,  
 CC fragments and conjugates are used to deplete levels of complement  
 CC pathway proteins (by superactivation until one or more components are  
 CC exhausted), specifically to prevent rejection of foreign material  
 CC (particularly a xenograft) but also to prevent complement-mediated  
 CC diseases resulting from (surgical) injury or antibody-antigen interaction  
 CC in autoimmune disease, also to localise and/or amplify endogenous  
 CC complement protein conversion and deposition at a specific site (e.g. a  
 CC virus, infected cell or tumour, to increase sensitivity to  
 CC complement-mediated responses; a particular application is eliminating  
 CC any cancer cells left after surgical removal of a tumour). Also  
 CC contemplated is ex vivo treatment, especially by passing blood through a  
 CC matrix containing (A) (this may remove additional anaphylactic peptides  
 CC and other inflammatory mediators) or killing of leukaemia cells or  
 CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
 CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
 CC inactivated), causing inactivation of the alternative pathway by  
 CC consumption of factor B.

CC Sequence 1663 AA;  
 SQ

Query Match 100.0%; Score 70; DB 18; Length 1663;  
 Best Local Similarity 100.0%; Pred. No. 0.0069;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHWESASLT 13  
 Db 1307 ITHRIHWESASLT 1319

RESULT 20  
 AAMW0988  
 ID AAMW0988 standard; Protein; 1663 AA.  
 AC AAMW0988;  
 XX  
 DT 09-APR-1998 (first entry)  
 DE Human C3 protein mutant R1303X, R1320X.  
 XX  
 KW Human; C3 protein; convertase; complement pathway protein; infection;  
 KW down-regulation resistant C3 convertase; xenograft rejection; therapy;  
 KW complement-mediated disease; autoimmune disease; leukaemia cell; tumour;  
 KW complement-mediated response; MHC-mismatched lymphocyte; murein.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1303  
 FT /label= Glu, Tyr, Cys, Trp, Gln, Gly  
 FT Misc-difference 1320  
 FT /label= Glu, Tyr, Cys, Trp, Gln, Gly  
 FT  
 XX W09732981-A1.  
 PN 12-SEP-1997.  
 XX  
 PD 04-MAR-1997; 97WO-GB00603.  
 XX  
 PF 19-NOV-1996; 96GB-0024028.  
 PR 07-MAR-1996; 96GB-0004865.  
 PR 07-JUN-1996; 96GB-0011896.  
 PR 08-JUL-1996; 96GB-0014293.  
 XX  
 PA (IMUT-) IMUTRAN LTD.  
 PA Farries TC, Harrison RA;  
 PI  
 XX

DR WPI; 1997-457534/42.

XX Modified complement pathway protein that forms C3 convertase  
 FT resistant to down-regulation - used to exhaust the complement  
 FT pathway by super-activation, especially for preventing graft  
 FT rejection, etc.  
 PR  
 XX  
 PS Claim 6; Page -; 123pp; English.

CC This sequence represents a mutated human C3 protein of the invention  
 CC (see AAMW3606 for wild type protein). This protein is a protein of the  
 CC invention, and is a modified native complement pathway protein (A) that  
 CC forms a down-regulation resistant C3 convertase. (A), their variants,  
 CC fragments and conjugates are used to deplete levels of complement  
 CC pathway proteins (by superactivation until one or more components are  
 CC exhausted), specifically to prevent rejection of foreign material  
 CC (particularly a xenograft) but also to prevent complement-mediated  
 CC diseases resulting from (surgical) injury or antibody-antigen interaction  
 CC in autoimmune disease, also to localise and/or amplify endogenous  
 CC complement protein conversion and deposition at a specific site (e.g. a  
 CC virus, infected cell or tumour, to increase sensitivity to  
 CC complement-mediated responses; a particular application is eliminating  
 CC any cancer cells left after surgical removal of a tumour). Also  
 CC contemplated is ex vivo treatment, especially by passing blood through a  
 CC matrix containing (A) (this may remove additional anaphylactic peptides  
 CC and other inflammatory mediators) or killing of leukaemia cells or  
 CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
 CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
 CC inactivated), causing inactivation of the alternative pathway by  
 CC consumption of factor B.

CC Sequence 1663 AA;  
 SQ

Query Match 100.0%; Score 70; DB 18; Length 1663;  
 Best Local Similarity 100.0%; Pred. No. 0.0069;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHWESASLT 13  
 Db 1307 ITHRIHWESASLT 1319

RESULT 21  
 AAMW0989  
 ID AAMW0989 standard; Protein; 1663 AA.  
 AC AAMW0989;  
 XX  
 DT 09-APR-1998 (first entry)  
 DE Human C3 protein mutant R1303X.  
 XX  
 KW Human; C3 protein; convertase; complement pathway protein; infection;  
 KW down-regulation resistant C3 convertase; xenograft rejection; therapy;  
 KW complement-mediated disease; autoimmune disease; leukaemia cell; tumour;  
 KW complement-mediated response; MHC-mismatched lymphocyte; murein.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1303  
 FT /label= Glu, Gln, Gly  
 FT  
 XX W09732981-A1.  
 PN 12-SEP-1997.  
 XX  
 PD 04-MAR-1997; 97WO-GB00603.  
 XX  
 PF 19-NOV-1996; 96GB-0024028.  
 PR 07-MAR-1996; 96GB-0004865.  
 PR 07-JUN-1996; 96GB-0011896.  
 PR 08-JUL-1996; 96GB-0014293.  
 XX

```

XX (IMOT-) IMUTRAN LTD.
PA
XX Farries TC, Harrison RA,
XX WPI; 1997-457534/42.
XX
XX Modified complement pathway protein that forms C3 convertase
XX PT resistant to down-regulation - used to exhaust the complement
XX PT pathway by super-activation, especially for preventing graft
XX PT rejection, etc.
XX
XX Claim 8; Page -: 123pp; English.
XX
XX This sequence represents a mutated human C3 protein of the invention
XX (see AAM34606 for wild type protein). This protein is a protein of the
XX invention, and is a modified native complement pathway protein (A) that
XX forms a down-regulation resistant C3 convertase. (A), their variants,
XX fragments and conjugates are used to deplete levels of complement
XX pathway proteins (by superactivation until one or more components are
XX exhausted), specifically to prevent rejection of foreign material
XX (particularly a xenograft) but also to prevent complement-mediated
XX diseases resulting from (surgical) injury or antibody-antigen interaction
XX in autoimmune disease, also to localise and/or amplify endogenous
XX complement protein conversion and deposition at a specific site (e.g. a
XX virus, infected cell or tumour, to increase sensitivity to
XX complement-mediated responses; a particular application is eliminating
XX any cancer cells left after surgical removal of a tumour). Also
XX contemplated is ex vivo treatment, especially by passing blood through a
XX matrix containing (A) (this may remove additional anaphylactic peptides
XX and other inflammatory mediators) or killing of leukaemia cells or
XX MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not
XX inhibited by factor I, it can bind repeatedly to factor B (which is then
XX inactivated), causing inactivation of the alternative pathway by
XX consumption of factor B.
XX
XX Sequence 1663 AA;
XX
XX Query Match 100.0%; Score 70; DB 18; Length 1663;
XX Best Local Similarity 100.0%; Pred. No. 0.0069;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ITHRIHWSASLL 13
XX |||||
XX Db 1307 ITHRIHWSASLL 1319
XX
XX RESULT 22
XX AAM40990
XX ID AAM40990 standard; Protein; 1663 AA.
XX AC AAM40990;
XX DT 09-APR-1998 (first entry)
XX XX Human C3 protein mutant.
XX DE Human C3 protein mutant.
XX XX Human; C3 protein; convertase; complement pathway protein; infection;
XX KM down-regulation resistant C3 convertase; xenograft rejection; therapy;
XX KM complement-mediated disease; autoimmune disease; leukaemia cell; tumour;
XX KM complement-mediated response; MHC-mismatched lymphocyte; mutein.
XX KW
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 992 /note= "E992S mutation"
XX FT Misc-difference 993 /note= "D993A mutation"
XX FT Misc-difference 996 /note= "D996S mutation"
XX FT Misc-difference 997 /note= "A997Q mutation"
XX FT

```

```

FT Misc-difference 998 /note= "E998S mutation"
FT FT Misc-difference 999 /note= "R999G mutation"
FT FT Misc-difference 1000 /note= "L1000M mutation"
FT FT Misc-difference 1001 /note= "K1001N mutation"
FT FT Misc-difference 1002 /note= "H1002I mutation"
FT FT Misc-difference 1005 /note= "V1005H mutation"
FT
FT WO9732981-A1.
FT
FT 12-SEP-1997.
FT
FT 04-MAR-1997; 97WO-GB00603.
FT
FT 19-NOV-1996; 96GB-0024028.
FT PR 07-MAR-1996; 96GB-0004865.
FT PR 07-JUN-1996; 96GB-0011896.
FT PR 08-JUL-1996; 96GB-0014293.
FT
FT (IMOT-) IMUTRAN LTD.
FT
FT Farries TC, Harrison RA;
FT WPI; 1997-457534/42.
FT
FT Modified complement pathway protein that forms C3 convertase
FT PT resistant to down-regulation - used to exhaust the complement
FT PT pathway by super-activation, especially for preventing graft
FT PT rejection, etc.
FT
FT Claim 15; Page -: 123pp; English.
FT
FT This sequence represents a mutated human C3 protein of the invention
FT (see AAM34606 for wild type protein). This protein is a protein of the
FT invention, and is a modified native complement pathway protein (A) that
FT forms a down-regulation resistant C3 convertase. (A), their variants,
FT fragments and conjugates are used to deplete levels of complement
FT pathway proteins (by superactivation until one or more components are
FT exhausted), specifically to prevent rejection of foreign material
FT (particularly a xenograft) but also to prevent complement-mediated
FT diseases resulting from (surgical) injury or antibody-antigen interaction
FT in autoimmune disease, also to localise and/or amplify endogenous
FT complement protein conversion and deposition at a specific site (e.g. a
FT virus, infected cell or tumour, to increase sensitivity to
FT complement-mediated responses; a particular application is eliminating
FT any cancer cells left after surgical removal of a tumour). Also
FT contemplated is ex vivo treatment, especially by passing blood through a
FT matrix containing (A) (this may remove additional anaphylactic peptides
FT and other inflammatory mediators) or killing of leukaemia cells or
FT MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not
FT inhibited by factor I, it can bind repeatedly to factor B (which is then
FT inactivated), causing inactivation of the alternative pathway by
FT consumption of factor B.
FT
FT SQ Sequence 1663 AA;
FT
FT Query Match 100.0%; Score 70; DB 18; Length 1663;
FT Best Local Similarity 100.0%; Pred. No. 0.0069;
FT Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT
FT QY 1 ITHRIHWSASLL 13
FT |||||
FT Db 1307 ITHRIHWSASLL 1319
FT
FT RESULT 23
FT AAM34606
FT ID AAM34606 standard; Protein; 1663 AA.

```



```

XX AC AAW34606;
XX DT 09-APR-1998 (first entry)
XX DE Wild type human C3 protein.
XX KM Human; C3 protein; convertase; complement pathway protein; infection;
XX KM down-regulation resistant C3 convertase; xenograft rejection; therapy;
XX KM complement-mediated disease; autoimmune disease; leukemia cell; tumour;
XX KM complement-mediated response; MHC-mismatched lymphocyte;
XX OS Homo sapiens.
XX PN MO9732981-A1.
XX PD 12-SEP-1997.
XX PF 04-MAR-1997; 97WO-GB00603.
XX PR 19-NOV-1996; 96GB-0024028.
XX PR 07-MAR-1996; 96GB-0004865.
XX PR 07-JUN-1996; 96GB-0011896.
XX PR 08-JUL-1996; 96GB-0014293.
XX PA (IMUT-) IMUTRAN LTD.
XX PI Farries TC, Harrison RA;
XX DR WPI; 1997-457534/42.
XX DR N-PSDB; AAT93749.
XX PT Modified complement pathway protein that forms C3 convertase
XX PT resistant to down-regulation - used to exhaust the complement
XX PT pathway by super-activation, especially for preventing graft
XX PT rejection, etc.
XX PS Disclosure; Fig 1; 123pp; English.
XX CC This sequence represents the wild type human C3 protein. This protein can
XX CC be mutated to produce a protein of the invention. The protein of the
XX CC invention is a modified native complement pathway protein (A) that forms
XX CC a down-regulation resistant C3 convertase. (A), their variants, fragments
XX CC and conjugates are used to deplete levels of complement pathway proteins
XX CC (by superactivation until one or more components are exhausted),
XX CC specifically to prevent rejection of foreign material (particularly a
XX CC xenograft) but also to prevent complement-mediated diseases resulting
XX CC from (surgical) injury or antibody-antigen interaction in autoimmune
XX CC disease, also to localise and/or amplify endogenous complement protein
XX CC conversion and deposition at a specific site (e.g. a virus, infected cell
XX CC or tumour, to increase sensitivity to complement-mediated responses; a
XX CC particular application is eliminating any cancer cells left after
XX CC surgical removal of a tumour). Also contemplated is ex vivo treatment,
XX CC especially by passing blood through a matrix containing (A) (this may
XX CC remove additional anaphylactic peptides and other inflammatory
XX CC mediators) or killing of leukemia cells or MHC-mismatched lymphocytes in
XX CC extracted bone marrow. Since (A) is not inhibited by factor I, it can
XX CC bind repeatedly to factor B (which is then inactivated), causing
XX CC inactivation of the alternative pathway by consumption of factor B.
XX SQ Sequence 1663 AA;
XX
Query Match 100.0%; Score 70; DB 16; Length 1663;
Best Local Similarity 100.0%; Pred. No. 0.0069;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

ID ID AAW34607 standard; Protein; 1663 AA.
XX AC AAW34607;
XX DT 09-APR-1998 (first entry)
XX DE Human C3 protein mutant C3M-123.
XX KM Human; C3 protein; convertase; complement pathway protein; infection;
XX KM down-regulation resistant C3 convertase; xenograft rejection; therapy;
XX KM complement-mediated disease; autoimmune disease; leukemia cell; tumour;
XX KM complement-mediated response; MHC-mismatched lymphocyte; mutin.
XX OS Homo sapiens.
XX PN MO9732981-A1.
XX PD 12-SEP-1997.
XX PF 04-MAR-1997; 97WO-GB00603.
XX PR 19-NOV-1996; 96GB-0024028.
XX PR 07-MAR-1996; 96GB-0004865.
XX PR 07-JUN-1996; 96GB-0011896.
XX PR 08-JUL-1996; 96GB-0014293.
XX PA (IMUT-) IMUTRAN LTD.
XX PI Farries TC, Harrison RA;
XX DR WPI; 1997-457534/42.
XX PT Modified complement pathway protein that forms C3 convertase
XX PT resistant to down-regulation - used to exhaust the complement
XX PT pathway by super-activation, especially for preventing graft
XX PT rejection, etc.
XX PS Example 4; Page -; 123pp; English.
XX CC This sequence represents a mutated human C3 protein of the invention
XX CC (see AAW34606 for wild type protein). This protein is a protein of the
XX CC invention, and is a modified native complement pathway protein (A) that
XX CC forms a down-regulation resistant C3 convertase. (A), their variants,
XX CC fragments and conjugates are used to deplete levels of complement
XX CC pathway proteins (by superactivation until one or more components are
XX CC exhausted), specifically to prevent rejection of foreign material
XX CC (particularly a xenograft) but also to prevent complement-mediated
XX CC diseases resulting from (surgical) injury or antibody-antigen interaction
XX CC in autoimmune disease, also to localise and/or amplify endogenous
XX CC complement protein conversion and deposition at a specific site (e.g. a
XX CC virus, infected cell or tumour, to increase sensitivity to
XX CC complement-mediated responses; a particular application is eliminating
XX CC any cancer cells left after surgical removal of a tumour). Also
XX CC contemplated is ex vivo treatment, especially by passing blood through a
XX CC matrix containing (A) (this may remove additional anaphylactic peptides
XX CC and other inflammatory mediators) or killing of leukemia cells or
XX CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not
XX CC inhibited by factor I, it can bind repeatedly to factor B (which is then
XX CC inactivated), causing inactivation of the alternative pathway by
XX CC consumption of factor B.
XX SQ Sequence 1663 AA;
XX
Query Match 100.0%; Score 70; DB 16; Length 1663;
Best Local Similarity 100.0%; Pred. No. 0.0069;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 1307 ITHRIMESASL 1319

RESULT 25

AAW34608 standard; Protein; 1663 AA.

AC AAW34608;

DT 09-APR-1998 (first entry)

DE Human C3 protein mutant C3M-26.

KW Human; C3 protein; convertase; complement pathway protein; infection; down-regulation resistant C3 convertase; xenograft rejection; therapy; complement-mediated disease; autoimmune disease; leukaemia cell; tumour; complement-mediated response; MHC-mismatched lymphocyte; mutein.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 1303 /note= "Arg to Gln mutation"

FT Misc-difference 1320 /note= "Arg to Gln mutation"

PN WO9732981-A1.

XX 12-SEP-1997.

XX 04-MAR-1997; 97WO-GB00603.

XX 19-NOV-1996; 96GB-0024028.

PR 07-MAR-1996; 96GB-0004865.

PR 07-JUN-1996; 96GB-0011896.

PR 08-JUL-1996; 96GB-0014293.

XX (IMUT-) IMOTRAN LTD.

XX Farries TC, Harrison RA;

DR WPI; 1997-457534/42.

XX Modified complement pathway protein that forms C3 convertase resistant to down-regulation - used to exhaust the complement pathway by super-activation, especially for preventing graft rejection, etc.

XX Example 4; Page -; 123pp; English.

XX This sequence represents a mutated human C3 protein of the invention (see AAW34606 for wild type protein). This protein is a protein of the invention, and is a modified native complement pathway protein (A) that forms a down-regulation resistant C3 convertase. (A), their variants, fragments and conjugates are used to deplete levels of complement pathway proteins (by superactivation until one or more components are exhausted), specifically to prevent rejection of foreign material (particularly a xenograft) but also to prevent complement-mediated diseases resulting from (surgical) injury or antibody-antigen interaction in autoimmune disease, also to localise and/or amplify endogenous complement protein conversion and deposition at a specific site (e.g. a virus, infected cell or tumour, to increase sensitivity to complement-mediated responses; a particular application is eliminating any cancer cells left after surgical removal of a tumour). Also contemplated is ex vivo treatment, especially by passing blood through a matrix containing (A) (this may remove additional anaphylactic peptides and other inflammatory mediators) or killing of leukaemia cells or MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not inhibited by factor I, it can bind repeatedly to factor B (which is then inactivated), causing inactivation of the alternative pathway by consumption of factor B.

XX Sequence 1663 AA;

Query Match 100.0%; Score 70; DB 18; Length 1663;

Best Local Similarity 100.0%; Pred. No. 0.0069;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITHRIMESASL 13

Db 1307 ITHRIMESASL 1319

RESULT 26

AAW34609 standard; Protein; 1663 AA.

AC AAW34609;

DT 09-APR-1998 (first entry)

DE Human C3 protein mutant C3M-51.

KW Human; C3 protein; convertase; complement pathway protein; infection; down-regulation resistant C3 convertase; xenograft rejection; therapy; complement-mediated disease; autoimmune disease; leukaemia cell; tumour; complement-mediated response; MHC-mismatched lymphocyte; mutein.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 1320 /note= "Arg to Gln mutation"

PN WO9732981-A1.

XX 12-SEP-1997.

XX 04-MAR-1997; 97WO-GB00603.

XX 19-NOV-1996; 96GB-0024028.

PR 07-MAR-1996; 96GB-0004865.

PR 07-JUN-1996; 96GB-0011896.

PR 08-JUL-1996; 96GB-0014293.

XX (IMUT-) IMOTRAN LTD.

XX Farries TC, Harrison RA;

DR WPI; 1997-457534/42.

XX Modified complement pathway protein that forms C3 convertase resistant to down-regulation - used to exhaust the complement pathway by super-activation, especially for preventing graft rejection, etc.

XX Claim 7; Page -; 123pp; English.

XX This sequence represents a mutated human C3 protein of the invention (see AAW34606 for wild type protein). This protein is a protein of the invention, and is a modified native complement pathway protein (A) that forms a down-regulation resistant C3 convertase. (A), their variants, fragments and conjugates are used to deplete levels of complement pathway proteins (by superactivation until one or more components are exhausted), specifically to prevent rejection of foreign material (particularly a xenograft) but also to prevent complement-mediated diseases resulting from (surgical) injury or antibody-antigen interaction in autoimmune disease, also to localise and/or amplify endogenous complement protein conversion and deposition at a specific site (e.g. a virus, infected cell or tumour, to increase sensitivity to complement-mediated responses; a particular application is eliminating any cancer cells left after surgical removal of a tumour). Also contemplated is ex vivo treatment, especially by passing blood through a matrix containing (A) (this may remove additional anaphylactic peptides and other inflammatory mediators) or killing of leukaemia cells or MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not



CC invention, and is a modified native complement pathway protein (A) that  
 CC forms a down-regulation resistant C3 convertase. (A), their variants,  
 CC fragments and conjugates are used to deplete levels of complement  
 CC pathway proteins (by superactivation until one or more components are  
 CC exhausted), specifically to prevent rejection of foreign material  
 CC (particularly a xenograft) but also to prevent complement-mediated  
 CC diseases resulting from (surgical) injury or antibody-antigen interaction  
 CC in autoimmune disease, also to localise and/or amplify endogenous  
 CC complement protein conversion and deposition at a specific site (e.g. a  
 CC virus, infected cell or tumour, to increase sensitivity to  
 CC complement-mediated responses; a particular application is eliminating  
 CC any cancer cells left after surgical removal of a tumour). Also  
 CC contemplated is ex vivo treatment, especially by passing blood through a  
 CC matrix containing (A) (this may remove additional anaphylactic peptides  
 CC and other inflammatory mediators) or killing of leukaemia cells or  
 CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
 CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
 CC inactivated), causing inactivation of the alternative pathway by  
 CC consumption of factor B.

CC Sequence 1663 AA;

CC Query Match 100.0%; Score 70; DB 18; Length 1663;

CC Best Local Similarity 100.0%; Pred. No. 0.0069;

CC Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ITHRIHWSASLL 13

DB 1307 ITHRIHWSASLL 1319

RESULT 29

AAW34612

ID AAW34612 standard; Protein; 1663 AA.

AC AAW34612;

DT 09-APR-1998 (first entry)

DE Human C3 protein mutant DV-1AM.

KM Human; C3 protein; convertase; complement pathway protein; infection;

KW down-regulation resistant C3 convertase; xenograft rejection; therapy;

KM complement-mediated disease; autoimmune disease; leukaemia cell; tumour;

KW complement-mediated response; MHC-mismatched lymphocyte; mutein.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 992 /note= "E992S mutation"

FT Misc-difference 993 /note= "D993A mutation"

FT Misc-difference 996 /note= "D996S mutation"

FT Misc-difference 997 /note= "D996S mutation"

FT Misc-difference 998 /note= "A997Q mutation"

FT Misc-difference 999 /note= "E996S mutation"

FT Misc-difference 1000 /note= "R999G mutation"

FT Misc-difference 1000 /note= "L1000M mutation"

PN WO9732981-A1.

PD 12-SEP-1997.

PF 04-MAR-1997; 97WO-GB00603.

XX 19-NOV-1996; 96GB-0024028.

XX 07-MAR-1996; 96GB-0004865.

PR 07-JUN-1996; 96GB-0011896.

PR 08-JUL-1996; 96GB-0014293.

XX (IMUT-) IMUTRAN LTD.

PI Farries TC, Harrison RA;

DR WPI; 1997-457534/42.

PT Modified complement pathway protein that forms C3 convertase  
 PT resistant to down-regulation - used to exhaust the complement  
 PT pathway by super-activation, especially for preventing graft  
 PT rejection, etc.

PS Example 14; Page -; 123pp; English.

XX This sequence represents a mutated human C3 protein of the invention  
 CC (see AAW34606 for wild type protein). This protein is a protein of the  
 CC invention, and is a modified native complement pathway protein (A) that  
 CC forms a down-regulation resistant C3 convertase. (A), their variants,  
 CC fragments and conjugates are used to deplete levels of complement  
 CC pathway proteins (by superactivation until one or more components are  
 CC exhausted), specifically to prevent rejection of foreign material  
 CC (particularly a xenograft) but also to prevent complement-mediated  
 CC diseases resulting from (surgical) injury or antibody-antigen interaction  
 CC in autoimmune disease, also to localise and/or amplify endogenous  
 CC complement protein conversion and deposition at a specific site (e.g. a  
 CC virus, infected cell or tumour, to increase sensitivity to  
 CC complement-mediated responses; a particular application is eliminating  
 CC any cancer cells left after surgical removal of a tumour). Also  
 CC contemplated is ex vivo treatment, especially by passing blood through a  
 CC matrix containing (A) (this may remove additional anaphylactic peptides  
 CC and other inflammatory mediators) or killing of leukaemia cells or  
 CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
 CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
 CC inactivated), causing inactivation of the alternative pathway by  
 CC consumption of factor B.

SQ Sequence 1663 AA;

CC Query Match 100.0%; Score 70; DB 18; Length 1663;

CC Best Local Similarity 100.0%; Pred. No. 0.0069;

CC Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ITHRIHWSASLL 13

DB 1307 ITHRIHWSASLL 1319

RESULT 30

AAW34613

ID AAW34613 standard; Protein; 1663 AA.

AC AAW34613;

DT 09-APR-1998 (first entry)

DE Human C3 protein mutant DV-1B.

KM Human; C3 protein; convertase; complement pathway protein; infection;

KW down-regulation resistant C3 convertase; xenograft rejection; therapy;

KM complement-mediated disease; autoimmune disease; leukaemia cell; tumour;

KW complement-mediated response; MHC-mismatched lymphocyte; mutein.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 1001 /note= "K1001N mutation"

FT Misc-difference 1002 /note= "H1002I mutation"

FT Misc-difference 1005 /note= "V1005H mutation"

PN W09732981-A1.  
 XX 12-SEP-1997.  
 PD 04-MAR-1997; 97WO-GB00603.  
 XX 19-NOV-1996; 96GB-0024028.  
 PR 07-MAR-1996; 96GB-0004865.  
 PR 07-JUN-1996; 96GB-0011896.  
 PR 08-JUL-1996; 96GB-0014293.  
 XX (IMUT-) IMUTRAN LTD.  
 PA Farries TC, Harrison RA;  
 PI WPI; 1997-457534/42.  
 DR  
 XX Modified complement pathway protein that forms C3 convertase  
 PT resistant to down-regulation - used to exhaust the complement  
 PT pathway by super-activation, especially for preventing graft  
 PT rejection, etc.  
 XX  
 PS Example 14; Page -; 123p; English.  
 XX  
 CC This sequence represents a mutated human C3 protein of the invention  
 CC (see AAW34606 for wild type protein). This protein is a protein of the  
 CC invention, and is a modified native complement pathway protein (A) that  
 CC forms a down-regulation resistant C3 convertase. (A), their variants,  
 CC fragments and conjugates are used to deplete levels of complement  
 CC pathway proteins (by superactivation until one or more components are  
 CC exhausted), specifically to prevent rejection of foreign material  
 CC (particularly a xenograft) but also to prevent complement-mediated  
 CC diseases resulting from (surgical) injury or antibody-antigen interaction  
 CC in autoimmune disease, also to localise and/or amplify endogenous  
 CC complement protein conversion and deposition at a specific site (e.g. a  
 CC virus, infected cell or tumour, to increase sensitivity to  
 CC complement-mediated responses; a particular application is eliminating  
 CC any cancer cells left after surgical removal of a tumour). Also  
 CC contemplated is ex vivo treatment, especially by passing blood through a  
 CC matrix containing (A) (this may remove additional anaphylactic peptides  
 CC and other inflammatory mediators) or killing of leukaemia cells or  
 CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
 CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
 CC inactivated), causing inactivation of the alternative pathway by  
 CC consumption of factor B.  
 CC  
 XX Sequence 1663 AA;  
 SQ  
 Query Match 100.0%; Score 70; DB 18; Length 1663;  
 Best Local Similarity 100.0%; Pred. No. 0.0069;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ITHRIHWESASL 13  
 Db 1307 ITHRIHWESASL 1319  
 RESULT 31  
 AAW34614  
 ID AAW34614 standard; Protein; 1663 AA.  
 XX  
 AC AAW34614;  
 XX  
 DT 09-APR-1998 (first entry)  
 XX  
 DE Human C3 protein mutant DV-3.  
 XX  
 KM Human; C3 protein; convertase; complement pathway protein; infection;  
 KM down-regulation resistant C3 convertase; xenograft rejection; therapy;  
 KM complement-mediated disease; autoimmune disease; leukaemia cell; tumour;  
 KM complement-mediated response; MHC-mismatched lymphocyte; mteuin.  
 XX  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Misc-difference 1031 /note= "T1031G mutation"  
 FT Misc-difference 1032 /note= "E1032N mutation"  
 FT Misc-difference 1033 /note= "Q1033H mutation"  
 FT Misc-difference 1035 /note= "E1035N mutation"  
 FT Misc-difference 1036 /note= "K1036I mutation"  
 XX  
 PN W09732981-A1.  
 XX 12-SEP-1997.  
 PD 04-MAR-1997; 97WO-GB00603.  
 XX 19-NOV-1996; 96GB-0024028.  
 PR 07-MAR-1996; 96GB-0004865.  
 PR 07-JUN-1996; 96GB-0011896.  
 PR 08-JUL-1996; 96GB-0014293.  
 XX (IMUT-) IMUTRAN LTD.  
 PA Farries TC, Harrison RA;  
 PI WPI; 1997-457534/42.  
 DR  
 XX Modified complement pathway protein that forms C3 convertase  
 PT resistant to down-regulation - used to exhaust the complement  
 PT pathway by super-activation, especially for preventing graft  
 PT rejection, etc.  
 XX  
 PS Example 14; Page -; 123p; English.  
 XX  
 CC This sequence represents a mutated human C3 protein of the invention  
 CC (see AAW34606 for wild type protein). This protein is a protein of the  
 CC invention, and is a modified native complement pathway protein (A) that  
 CC forms a down-regulation resistant C3 convertase. (A), their variants,  
 CC fragments and conjugates are used to deplete levels of complement  
 CC pathway proteins (by superactivation until one or more components are  
 CC exhausted), specifically to prevent rejection of foreign material  
 CC (particularly a xenograft) but also to prevent complement-mediated  
 CC diseases resulting from (surgical) injury or antibody-antigen interaction  
 CC in autoimmune disease, also to localise and/or amplify endogenous  
 CC complement protein conversion and deposition at a specific site (e.g. a  
 CC virus, infected cell or tumour, to increase sensitivity to  
 CC complement-mediated responses; a particular application is eliminating  
 CC any cancer cells left after surgical removal of a tumour). Also  
 CC contemplated is ex vivo treatment, especially by passing blood through a  
 CC matrix containing (A) (this may remove additional anaphylactic peptides  
 CC and other inflammatory mediators) or killing of leukaemia cells or  
 CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
 CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
 CC inactivated), causing inactivation of the alternative pathway by  
 CC consumption of factor B.  
 CC  
 XX Sequence 1663 AA;  
 SQ  
 Query Match 100.0%; Score 70; DB 18; Length 1663;  
 Best Local Similarity 100.0%; Pred. No. 0.0069;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ITHRIHWESASL 13  
 Db 1307 ITHRIHWESASL 1319  
 RESULT 32  
 AAW34615  
 ID AAW34615 standard; Protein; 1663 AA.

XX	AAW34615;	
XX		
DT	09-APR-1998	(first entry)
XX		
DE	Human C3 protein mutant DV-4.	
XX		
XX	Human; C3 protein; convertase; complement pathway protein; infection;	
KM	down-regulation resistant C3 convertase; xenograft rejection; therapy;	
KM	complement-mediated disease; autoimmune disease; leukaemia cell; tumour;	
XX	complement-mediated response; MHC-mismatched lymphocyte; mteirn.	
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference	1070
FT	/note=	"V1070K mutation"
FT	Misc-difference	1071
FT	/note=	"K1071G mutation"
FT	Misc-difference	1072
FT	/note=	"R1072G mutation"
FT	Misc-difference	1073
FT	/note=	"A1073S mutation"
FT	Misc-difference	1074
FT	/note=	"P1074A mutation"
XX		
PN	WO9732381-A1.	
XX		
PD	12-SEP-1997.	
XX		
PF	04-MAR-1997;	97WO-GB00603.
XX		
PR	19-NOV-1996;	96GB-0024028.
PR	07-MAR-1996;	96GB-0004865.
PR	07-JUN-1996;	96GB-0011896.
PR	08-JUL-1996;	96GB-0014293.
XX		
PA	(IMUT-) IMUTRAN LTD.	
XX		
PI	Farries TC, Harrison RA;	
XX		
DR	WPI; 1997-457534/42.	
XX		
PT	Modified complement pathway protein that forms C3 convertase	
PT	resistant to down-regulation - used to exhaust the complement	
PT	pathway by super-activation, especially for preventing graft	
PT	rejection, etc.	
XX		
PS	Example 14; Page -; 123p; English.	
XX		
CC	This sequence represents a mutated human C3 protein of the invention	
CC	(see AAW34606 for wild type protein). This protein is a protein of the	
CC	invention, and is a modified native complement pathway protein (A) that	
CC	forms a down-regulation resistant C3 convertase. (A), their variants,	
CC	fragments and conjugates are used to deplete levels of complement	
CC	pathway proteins (by superactivation until one or more components are	
CC	exhausted), specifically to prevent rejection of foreign material	
CC	(particularly a xenograft) but also to prevent complement-mediated	
CC	diseases resulting from (surgical) injury or antibody-antigen interaction	
CC	in autoimmune disease, also to localise and/or amplify endogenous	
CC	complement protein conversion and deposition at a specific site (e.g. a	
CC	virus, infected cell or tumour, to increase sensitivity to	
CC	complement-mediated responses; a particular application is eliminating	
CC	any cancer cells left after surgical removal of a tumour). Also	
CC	contemplated is ex vivo treatment, especially by passing blood through a	
CC	matrix containing (A) (this may remove additional anaphylactic peptides	
CC	and other inflammatory mediators) or killing of leukaemia cells or	
CC	MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not	
CC	inhibited by factor I, it can bind repeatedly to factor B (which is then	
CC	inactivated), causing inactivation of the alternative pathway by	
CC	consumption of factor B.	
XX		
Sequence	1663 AA;	

Query Match	Similarity	100.0%	Score 70:	DB 18:	Length 1663;
Best Local	Similarity	100.0%	Pred. NO. 0.0069;		
Matches	13;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
CV	1	ITTHRIHWESASL 13			
DB	1307	ITTHRIHWESASL 1319			
RESULT 33					
AAW34616					
ID	AAW34616	standard; Protein; 1663	AA.		
XX					
AC	AAW34616;				
XX					
DT	09-APR-1998	(First entry)			
XX					
DE	Human C3 protein mutant CV-5.				
XX					
KM	Human; C3 protein; convertase; complement pathway protein; infection;				
KM	down-regulation resistant C3 convertase; xenograft rejection; therapy;				
KM	complement-mediated disease; autoimmune disease; leukaemia cell; tumour;				
XX	complement-mediated response; MHC-mismatched lymphocyte; mulein.				
XX					
OS	Homo sapiens.				
XX					
PH	Key	Location/Qualifiers			
FT	Misc-difference	1134			
XX		/note= "R1134Q mutation"			
PN	WO9732981-A1.				
XX					
PD	12-SEP-1997.				
XX					
PF	04-MAR-1997;	97WO-GB00603.			
XX					
PR	19-NOV-1996;	96GB-0024028.			
PR	07-MAR-1996;	96GB-0004865.			
PR	07-JUN-1996;	96GB-0011896.			
PR	08-JUL-1996;	96GB-0014293.			
XX					
PA	(IMUT-) IMUTRAN LTD.				
XX					
PI	Farries TC, Harrison RA;				
XX					
DR	WPI, 1997-457534/42.				
XX					
PT	Modified complement pathway protein that forms C3 convertase				
PT	resistant to down-regulation - used to exhaust the complement				
PT	pathway by super-activation, especially for preventing graft				
PT	rejection, etc.				
XX					
PS	Example 14; Page -; 123pp; English.				
XX					
CC	This sequence represents a mutated human C3 protein of the invention				
CC	(see AAW34606 for wild type protein). This protein is a protein of the				
CC	invention, and is a modified native complement pathway protein (A) that				
CC	forms a down-regulation resistant C3 convertase. (A), their variants,				
CC	fragments and conjugates are used to deplete levels of complement				
CC	pathway proteins (by superactivation until one or more components are				
CC	exhausted), specifically to prevent rejection of foreign material				
CC	(particularly a xenograft) but also to prevent complement-mediated				
CC	diseases resulting from (surgical) injury or antibody-antigen interaction				
CC	in autoimmune disease, also to localize and/or amplify endogenous				
CC	complement protein conversion and deposition at a specific site (e.g. a				
CC	virus, infected cell or tumour, to increase sensitivity to				
CC	complement-mediated responses; a particular application is eliminating				
CC	any cancer cells left after surgical removal of a tumour). Also				
CC	contemplated is ex vivo treatment, especially by passing blood through a				
CC	matrix containing (A) (this may remove additional anaphylactic peptides				
CC	and other inflammatory mediators) or killing of leukaemia cells or				
CC	MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not				

CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
CC inactivated), causing inactivation of the alternative pathway by  
CC consumption of factor B.  
XX  
SQ Sequence 1663 AA;  
Query Match 100.0%; Score 70; DB 18; Length 1663;  
Best Local Similarity 100.0%; Pred. No. 0.0069;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ITHRIHWSASL 13  
DB 1307 ITHRIHWSASL 1319  
RESULT 34  
AAW34617  
ID AAW34617 standard; Protein; 1663 AA.  
XX  
XX AAW34617;  
XX  
XX 09-APR-1998 (first entry)  
XX  
XX Human C3 protein mutant DV-6.  
XX  
XX Human; C3 protein; convertase; complement pathway protein; infection;  
XX down-regulation resistant C3 convertase; xenograft rejection; therapy;  
XX complement-mediated disease; autoimmune disease; leukaemia cell; tumour;  
XX complement-mediated response; MHC-mismatched lymphocyte; mutein.  
XX  
XX Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
XX FT Misc-difference 1152 /note= "Q1152R mutation"  
XX FT Misc-difference 1153 /note= "E1153K mutation"  
XX FT Misc-difference 1155 /note= "K1155F mutation"  
XX  
XX W09732981-A1.  
XX  
XX 12-SEP-1997.  
XX  
XX 04-MAR-1997; 97WO-GB00603.  
XX  
XX 19-NOV-1996; 96GB-0024028.  
XX 07-MAR-1996; 96GB-0004865.  
XX 07-JUN-1996; 96GB-0011896.  
XX 08-JUL-1996; 96GB-0014293.  
XX  
XX (IMUT-) IMUTRAN LTD.  
XX  
XX Farries TC, Harrison RA;  
XX  
XX WPI; 1997-457534/42.  
XX  
XX Modified complement pathway protein that forms C3 convertase  
XX resistant to down-regulation - used to exhaust the complement  
XX pathway by super-activation, especially for preventing graft  
XX rejection, etc.  
XX  
XX Claim 18; Page -; 123pp; English.  
XX  
XX This sequence represents a mutated human C3 protein of the invention  
XX (see AAW34606 for wild type protein). This protein is a protein of the  
XX invention, and is a modified native complement pathway protein (A) that  
XX forms a down-regulation resistant C3 convertase. (A), their variants,  
XX fragments and conjugates are used to deplete levels of complement  
XX pathway proteins (by superactivation until one or more components are  
XX exhausted), specifically to prevent rejection of foreign material  
XX (particularly a xenograft) but also to prevent complement-mediated  
XX diseases resulting from (surgical) injury or antibody-antigen interaction

CC in autoimmune disease, also to localise and/or amplify endogenous  
CC complement protein conversion and deposition at a specific site (e.g. a  
CC virus, infected cell or tumour, to increase sensitivity to  
CC complement-mediated responses; a particular application is eliminating  
CC any cancer cells left after surgical removal of a tumour). Also  
CC contemplated is ex vivo treatment, especially by passing blood through a  
CC matrix containing (A) (this may remove additional anaphylactic peptides  
CC and other inflammatory mediators) or killing of leukaemia cells or  
CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
CC inactivated), causing inactivation of the alternative pathway by  
CC consumption of factor B.  
XX  
XX SQ Sequence 1663 AA;  
Query Match 100.0%; Score 70; DB 18; Length 1663;  
Best Local Similarity 100.0%; Pred. No. 0.0069;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ITHRIHWSASL 13  
DB 1307 ITHRIHWSASL 1319  
RESULT 35  
AAW34618  
ID AAW34618 standard; Protein; 1663 AA.  
XX  
XX AAW34618;  
XX  
XX 09-APR-1998 (first entry)  
XX  
XX Human C3 protein mutant DV-7N.  
XX  
XX Human; C3 protein; convertase; complement pathway protein; infection;  
XX down-regulation resistant C3 convertase; xenograft rejection; therapy;  
XX complement-mediated disease; autoimmune disease; leukaemia cell; tumour;  
XX complement-mediated response; MHC-mismatched lymphocyte; mutein.  
XX  
XX Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
XX FT Misc-difference 1174 /note= "D1174N mutation"  
XX  
XX W09732981-A1.  
XX  
XX 12-SEP-1997.  
XX  
XX 04-MAR-1997; 97WO-GB00603.  
XX  
XX 19-NOV-1996; 96GB-0024028.  
XX 07-MAR-1996; 96GB-0004865.  
XX 07-JUN-1996; 96GB-0011896.  
XX 08-JUL-1996; 96GB-0014293.  
XX  
XX (IMUT-) IMUTRAN LTD.  
XX  
XX Farries TC, Harrison RA;  
XX  
XX WPI; 1997-457534/42.  
XX  
XX Modified complement pathway protein that forms C3 convertase  
XX resistant to down-regulation - used to exhaust the complement  
XX pathway by super-activation, especially for preventing graft  
XX rejection, etc.  
XX  
XX Example 14; Page -; 123pp; English.  
XX  
XX This sequence represents a mutated human C3 protein of the invention  
XX (see AAW34606 for wild type protein). This protein is a protein of the  
XX invention, and is a modified native complement pathway protein (A) that  
XX forms a down-regulation resistant C3 convertase. (A), their variants,

CC fragments and conjugates are used to deplete levels of complement  
CC pathway proteins (by superactivation until one or more components are  
CC exhausted), specifically to prevent rejection of foreign material  
CC (particularly a xenograft) but also to prevent complement-mediated  
CC diseases resulting from (surgical) injury or antibody-antigen interaction  
CC in autoimmune disease, also to localise and/or amplify endogenous  
CC complement protein conversion and deposition at a specific site (e.g. a  
CC virus, infected cell or tumour, to increase sensitivity to  
CC complement-mediated responses; a particular application is eliminating  
CC any cancer cells left after surgical removal of a tumour). Also  
CC contemplated is ex vivo treatment, especially by passing blood through a  
CC matrix containing (A) (this may remove additional anaphylactic peptides  
CC and other inflammatory mediators) or killing of leukaemia cells or  
CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
CC inactivated), causing inactivation of the alternative pathway by  
CC consumption of factor B.

CC Sequence 1663 AA;

Query Match 100.0%; Score 70; DB 18; Length 1663;

Best Local Similarity 100.0%; Pred. No. 0.0069; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHWSASLL 13

DB 1307 ITHRIHWSASLL 1319

RESULT 36

AAW34626 standard; Protein; 1667 AA.

AAW34626;

09-APR-1998 (first entry)

Human C3 protein mutant FT-4.

Human; C3 protein; convertase; complement pathway protein; infection;

down-regulation resistant C3 convertase; xenograft rejection; therapy;

complement-mediated disease; autoimmune disease; leukaemia cell; tumour;

complement-mediated response; MHC-mismatched lymphocyte; mutein.

Homo sapiens.

Key Location/Qualifiers

Misc-difference 1621..1625 /note="wild type residues IIGKD mutated to RYIYPLDSL"

WO9732981-A1.

12-SEP-1997.

04-MAR-1997; 97WO-GB00603.

19-NOV-1996; 96GB-0024028.

07-MAR-1996; 96GB-0004865.

07-JUN-1996; 96GB-0011896.

08-JUL-1996; 96GB-0014293.

(IMUT-) IMUTRAN LTD.

Parries TC, Harrison RA;

WPI; 1997-457534/42.

Modified complement pathway protein that forms C3 convertase

PT resistant to down-regulation - used to exhaust the complement

PT pathway by super-activation, especially for preventing graft

rejection, etc.

PS Example 17; Page -; 123pp; English.

XX This sequence represents a mutated human C3 protein of the invention  
CC (see AAW34606 for wild type protein). This protein is a protein of the  
CC invention, and is a modified native complement pathway protein (A) that  
CC forms a down-regulation resistant C3 convertase. (A), their variants,  
CC fragments and conjugates are used to deplete levels of complement  
CC pathway proteins (by superactivation until one or more components are  
CC exhausted), specifically to prevent rejection of foreign material  
CC (particularly a xenograft) but also to prevent complement-mediated  
CC diseases resulting from (surgical) injury or antibody-antigen interaction  
CC in autoimmune disease, also to localise and/or amplify endogenous  
CC complement protein conversion and deposition at a specific site (e.g. a  
CC virus, infected cell or tumour, to increase sensitivity to  
CC complement-mediated responses; a particular application is eliminating  
CC any cancer cells left after surgical removal of a tumour). Also  
CC contemplated is ex vivo treatment, especially by passing blood through a  
CC matrix containing (A) (this may remove additional anaphylactic peptides  
CC and other inflammatory mediators) or killing of leukaemia cells or  
CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
CC inactivated), causing inactivation of the alternative pathway by  
CC consumption of factor B.

CC Sequence 1667 AA;

Query Match 100.0%; Score 70; DB 18; Length 1667;

Best Local Similarity 100.0%; Pred. No. 0.0069; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHWSASLL 13

DB 1307 ITHRIHWSASLL 1319

RESULT 37

AAW34631 standard; Protein; 1667 AA.

AAW34631;

09-APR-1998 (first entry)

Human C3 protein mutant FR-4.

Human; C3 protein; convertase; complement pathway protein; infection;

down-regulation resistant C3 convertase; xenograft rejection; therapy;

complement-mediated disease; autoimmune disease; leukaemia cell; tumour;

complement-mediated response; MHC-mismatched lymphocyte; mutein.

Homo sapiens.

Key Location/Qualifiers

Misc-difference 1648 /note="wild type D mutated to AFLAN"

WO9732981-A1.

12-SEP-1997.

04-MAR-1997; 97WO-GB00603.

19-NOV-1996; 96GB-0024028.

07-MAR-1996; 96GB-0004865.

07-JUN-1996; 96GB-0011896.

08-JUL-1996; 96GB-0014293.

(IMUT-) IMUTRAN LTD.

Parries TC, Harrison RA;

WPI; 1997-457534/42.

Modified complement pathway protein that forms C3 convertase



PT resistant to down-regulation - used to exhaust the complement  
 PT pathway by super-activation, especially for preventing graft  
 rejection, etc.  
 XX  
 PS Example 17, Page -, 123pp; English.  
 XX  
 CC This sequence represents a mutated human C3 protein of the invention  
 CC (see AAM34606 for wild type protein). This protein is a protein of the  
 CC invention, and is a modified native complement pathway protein (A) that  
 CC forms a down-regulation resistant C3 convertase. (A), their variants,  
 CC fragments and conjugates are used to deplete levels of complement  
 CC pathway proteins (by superactivation until one or more components are  
 CC exhausted), specifically to prevent rejection of foreign material  
 CC (particularly a xenograft) but also to prevent complement-mediated  
 CC diseases resulting from (surgical) injury or antibody-antigen interaction  
 CC in autoimmune disease, also to localise and/or amplify endogenous  
 CC complement protein conversion and deposition at a specific site (e.g. a  
 CC virus, infected cell or tumour, to increase sensitivity to  
 CC complement-mediated responses; a particular application is eliminating  
 CC any cancer cells left after surgical removal of a tumour). Also  
 CC contemplated is ex vivo treatment, especially by passing blood through a  
 CC matrix containing (A) (this may remove additional anaphylactic peptides  
 CC and other inflammatory mediators) or killing of leukaemia cells or  
 CC MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not  
 CC inhibited by factor I, it can bind repeatedly to factor B (which is then  
 CC inactivated), causing inactivation of the alternative pathway by  
 CC consumption of factor B.  
 CC  
 XX  
 SQ Sequence 1667 AA;  
 Query Match 100.0%; Score 70; DB 18; Length 1667;  
 Best Local Similarity 100.0%; Pred. No. 0.0069;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ITHRIHWESASLL 13  
 ID |||||  
 Db 1307 ITHRIHWESASLL 1319  
 RESULT 38  
 AB08967  
 ID AB08967 standard; peptide; 11 AA.  
 XX  
 AC AB08967;  
 XX  
 DT 10-JUN-2003 (first entry)  
 XX  
 DE Human complement C3f derived peptide.  
 XX  
 KW Human complement; C3f; SERDI; myocardial infarction;  
 KW surface enhanced laser desorption/ionisation; intracerebral haemorrhage;  
 KW congestive heart failure; mass spectrometry; immunoassay;  
 KW radioimmunoassay; enzyme-linked immunosorbent assay; ELISA;  
 KW fluorescent immunoassay.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002161184-A1.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 30-APR-2001; 2001US-0845715.  
 XX  
 PR 30-APR-2001; 2001US-0845715.  
 XX  
 PA (JACK/) JACKOWSKI G.  
 PA (THAT/) THATCHER B.  
 PA (MARS/) MARSHALL J.  
 PA (YANT/) YANTHA J.  
 PA (VREE/) VREES T.  
 XX  
 PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;  
 XX

DR WPI; 2003-340874/32.  
 XX  
 XX Biopolymer marker, useful in diagnosing disease states including  
 PT myocardial infarction and intracerebral hemorrhage, comprises  
 PT complement C3f fragment with a specified molecular weight -  
 XX  
 PS Claim 1, Page 7, 11pp; English.  
 XX  
 CC The invention relates to a biopolymer marker comprises a complement C3f  
 CC fragment appearing as AB08967 having a molecular weight of 1348 daltons.  
 CC The fragment was isolated from samples from affected individuals using  
 CC SERDI (surface enhanced laser desorption/ionisation) and a protein chip.  
 CC The marker is useful in methods for diagnosing a disease state such as  
 CC myocardial infarction, intracerebral hemorrhage or congestive heart  
 CC failure. The methods used include mass spectroscopy or immunoassays, e.g.  
 CC radioimmunoassay, enzyme-linked immunosorbent assay (ELISA) or  
 CC fluorescent immunoassays. The invention enables a diagnostician to  
 CC characterise the presence or absence of the disease state relative to  
 CC recognition of the presence or absence of the biopolymer, respectively.  
 CC The present sequence is the human complement C3f derived peptide of  
 CC the invention.  
 CC  
 XX  
 SQ Sequence 11 AA;  
 Query Match 87.1%; Score 61; DB 24; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00083;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 HRIHWESASLL 13  
 ID |||||  
 Db 1 HRIHWESASLL 11  
 RESULT 39  
 ABG72971  
 ID ABG72971 standard; Peptide; 12 AA.  
 XX  
 AC ABG72971;  
 XX  
 DT 27-FEB-2003 (first entry)  
 XX  
 DE Human complement C3f fragment.  
 XX  
 KW Human biopolymer marker; myocardial infarction; C3f; syndrome X;  
 KW congestive heart failure; intracerebral haemorrhage; complement.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1 /note= "His is optionally absent"  
 FT Misc-difference 12 /note= "Arg is optionally absent"  
 FT  
 XX  
 PN WO200288731-A2.  
 XX  
 PD 07-NOV-2002.  
 XX  
 PF 29-APR-2002; 2002WO-CA00632.  
 XX  
 PR 30-APR-2001; 2001US-0845731.  
 XX  
 PA (SYNX-) SYN.X PHARMA INC.  
 XX  
 PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;  
 XX  
 DR WPI; 2003-111908/10.  
 XX  
 PT New biopolymer markers useful for indicating one particular disease  
 PT state such as myocardial infarction, congestive heart failure and  
 PT intracerebral hemorrhage -  
 XX  
 PS Claim 1, Page 30; 30pp; English.

XX The invention relates to a biopolymer marker having a sequence  
 CC useful in indicating at least one particular disease state. The  
 CC biopolymer marker is useful for indicating one particular disease state  
 CC such as syndrome X, myocardial infarction, congestive heart failure and  
 CC intracerebral haemorrhage. The marker is released into the circulation  
 CC where it may be present in the blood or any blood product. The presence  
 CC of each marker is determined using antibodies specific for each of the  
 CC markers and detecting specific binding of each antibody to its  
 CC respective marker. The present sequence represents a biopolymer marker  
 CC which is characterised as a human C3f fragment from the complement  
 CC system.  
 CC  
 SQ Sequence 12 AA;  
 Query Match 87.1%; Score 61; DB 24; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.00092;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 HRIHWESASL 13  
 |||||  
 1 HRIHWESASL 11

RESULT 40  
 ABB58617  
 ID ABB58617 standard; Protein; 411 AA.  
 XX  
 AC ABB58617;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 2643.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 XX pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 PI  
 PI Venter JC, Adams M, Li PMD, Myers EW;  
 XX  
 DR MPI; 2001-656860/75.  
 DR N-PSDB; ABL02720.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure; SEQ ID NO 2643; 21bp + Sequence listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX

SQ Sequence 411 AA;  
 Query Match 60.0%; Score 42; DB 22; Length 411;  
 Best Local Similarity 77.8%; Pred. No. 71;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 HRIHWESAS 11  
 ||: |||||  
 DB 150 HRIHWESAS 158  
 Search completed: August 28, 2003, 14:14:20  
 Job time : 85 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: August 28, 2003, 14:10:14 ; Search time 29 Seconds  
(without alignments)  
18.967 Million cell updates/sec

Title: US-09-845-738C-1  
Perfect score: 70  
Sequence: 1 ITHRIHWEASL 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/6C\_COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/6D\_COMB.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	70	100.0	1663	US-08-793-126-1	Sequence 1, Appli
2	70	100.0	1663	US-09-132-271-1	Sequence 1, Appli
3	70	100.0	1663	US-09-142-334-22	Sequence 22, Appli
4	46	65.7	281	US-09-252-991A-22644	Sequence 22644, A
5	42	60.0	280	US-09-252-991A-21635	Sequence 21635, A
6	39	55.7	221	US-09-252-991A-24616	Sequence 24616, A
7	39	55.7	264	US-09-107-532A-5290	Sequence 5290, Ap
8	39	55.7	329	US-09-107-532A-7038	Sequence 7038, Ap
9	37	52.9	267	US-09-449-218D-43	Sequence 43, Appli
10	37	52.9	267	US-09-668-529A-43	Sequence 43, Appli
11	37	52.9	267	US-09-668-037A-43	Sequence 43, Appli
12	37	52.9	272	US-08-887-997B-2	Sequence 2, Appli
13	37	52.9	526	US-08-311-731A-291	Sequence 291, App
14	37	52.9	538	US-08-311-731A-71	Sequence 71, Appli
15	37	52.9	700	US-09-413-814-68	Sequence 68, Appli
16	36	51.4	93	US-08-839-710-3	Sequence 3, Appli
17	36	51.4	93	US-09-066-262-3	Sequence 3, Appli
18	36	51.4	623	US-08-959-004-10	Sequence 10, Appli
19	36	51.4	844	US-07-646-537B-2	Sequence 2, Appli
20	35	50.0	233	US-09-214-631-7	Sequence 7, Appli
21	35	50.0	234	US-08-299-567-5	Sequence 5, Appli
22	35	50.0	238	US-08-240-124-2	Sequence 2, Appli
23	35	50.0	238	US-08-453-943-2	Sequence 2, Appli
24	35	50.0	238	US-09-057-121-2	Sequence 2, Appli
25	35	50.0	238	US-09-358-734-2	Sequence 2, Appli
26	35	50.0	332	US-09-134-001C-3977	Sequence 3977, Ap
27	35	50.0	356	US-09-252-991A-28288	Sequence 28288, A

28	35	50.0	542	US-09-252-991A-11968	Sequence 31968, A
29	35	50.0	556	US-08-501-572-1	Sequence 1, Appli
30	35	50.0	556	US-09-040-444-1	Sequence 1, Appli
31	35	50.0	566	US-09-491-522-7	Sequence 7, Appli
32	35	50.0	595	US-09-252-991A-18995	Sequence 18995, A
33	35	50.0	796	US-09-107-532A-7065	Sequence 7065, Ap
34	35	50.0	860	US-09-252-991A-26112	Sequence 26112, A
35	35	50.0	1205	US-09-491-532-11	Sequence 11, Appli
36	35	50.0	1211	US-09-491-532-5	Sequence 5, Appli
37	34	48.6	105	US-08-889-013C-6	Sequence 6, Appli
38	34	48.6	126	US-09-107-532A-3813	Sequence 3813, Ap
39	34	48.6	166	US-08-729-103-4	Sequence 4, Appli
40	34	48.6	226	US-09-252-991A-19372	Sequence 19372, A
41	34	48.6	260	US-09-252-991A-28140	Sequence 28140, A
42	34	48.6	279	US-09-252-991A-32422	Sequence 32422, A
43	34	48.6	288	US-08-632-434-8	Sequence 8, Appli
44	34	48.6	313	US-08-632-434-6	Sequence 6, Appli
45	34	48.6	320	US-08-632-434-2	Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-08-793-126-1  
; Sequence 1, Application US/08793126  
; Patent No. 5849297  
; GENERAL INFORMATION:  
; APPLICANT: Harrison, Richard Alexander  
; APPLICANT: Faries, Charles Timothy  
; TITLE OF INVENTION: MODIFIED HUMAN C3 PROTEINS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: HALE AND DORR LLP  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: United States of America  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/793,126  
; FILING DATE: 07-FEB-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baker, Hollie L.  
; REGISTRATION NUMBER: 31,321  
; TELEPHONE: (617) 526-6000  
; TELEPHONE: (617) 526-6000  
; TELEFAX: (617) 526-5000  
; INFORMATION FOR SEQ ID NO. 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1663 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-793-126-1

Query Match 100.0%; Score 70; DB 2; Length 1663;  
Best Local Similarity 100.0%; Pred. No. 0.0036;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 ITHRIHWEASL 13  
Db 1307 ITHRIHWEASL 1319

```

RESULT 2
US-09-132-271-1
Sequence 1, Application US/09132271
Patent No. 6221657
GENERAL INFORMATION:
APPLICANT: Harrison, Richard Alexander
APPLICANT: Fairies, Charles Timothy
TITLE OF INVENTION: MODIFIED HUMAN C3 PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: HALE AND DORR LLP
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: United States of America
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/132,271
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/793,126
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Baker, Hollie L.
REGISTRATION NUMBER: 31,321
REFERENCE/DOCKET NUMBER: 102286.377
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-5000
INFORMATION FOR SEQ ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1663 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-132-271-1

Query Match 100.0%; Score 70; DB 3; Length 1663;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

OY 1 ITTHRIHWESASLL 13
Db 1307 ITTHRIHWESASLL 1319

RESULT 3
US-09-142-334-22
Sequence 22, Application US/09142334
Patent No. 6268485
GENERAL INFORMATION:
APPLICANT: Fairies, Timothy C.
APPLICANT: Harrison, Richard A.
TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase
FILE REFERENCE: 4-30443/A/INT/PCF
CURRENT FILING DATE: 1999-04-15
CURRENT APPLICATION NUMBER: US/09/142,334
EARLIER APPLICATION NUMBER: PCT/GB97/00603
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 22
LENGTH: 1663
TYPE: PRT
ORGANISM: Homo sapiens
US-09-142-334-22

```

```

Query Match          100.0%; Score 70; DB 3; Length 1663;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ITRIHVESASL 13
|||||
Db      1307 ITRIHVESASL 1319

RESULT 4
US-09-252-991A-22644
; Sequence 22644, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22644
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22644

Query Match          65.7%; Score 46; DB 4; Length 281;
Best Local Similarity 58.3%; Pred. No. 4.2;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 ITRIHVESASL 12
|||:|
Db      134 IAHHLHWGHASL 145

RESULT 5
US-09-252-991A-21635
; Sequence 21635, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21635
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21635

Query Match          60.0%; Score 42; DB 4; Length 280;
Best Local Similarity 46.2%; Pred. No. 18;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      1 ITRIHVESASL 13
|||:|
Db      121 IVHRLDWTETGSLM 133

RESULT 6
```

US-09-252-991A-24616  
; Sequence 24616, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 24616  
; LENGTH: 221  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24616

Query Match 55.7%; Score 39; DB 4; Length 221;  
Best Local Similarity 66.7%; Pred. No. 43;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 2 THRHWESASLL 13  
Db 142 TPRHMSASML 153

RESULT 7  
US-09-107-532A-5290  
; Sequence 5290, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneka  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 5290:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 264 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES

ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (B) LOCATION 1...264  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5290:  
US-09-107-532A-5290

Query Match 55.7%; Score 39; DB 4; Length 264;  
Best Local Similarity 53.3%; Pred. No. 52;  
Matches 8; Conservative 4; Mismatches 1; Indels 2; Gaps 1;

Oy 1 ITHRI--HWESASLL 13  
Db 86 VTHRIHWMAALL 100

RESULT 8  
US-09-107-532A-7038  
; Sequence 7038, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneka  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 7038:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 329 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (B) LOCATION 1...329  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7038:  
US-09-107-532A-7038

Query Match 55.7%; Score 39; DB 4; Length 329;  
Best Local Similarity 58.3%; Pred. No. 65;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 2 THRHWESASLL 13

DB 105 THRVNFDLSALL 116

||||:||||

RESULT 9  
US-09-449-218D-43  
; Sequence 43, Application US/09449218D  
; Patent No. 6395511  
; GENERAL INFORMATION:  
; APPLICANT: Brunkow, Mary E.  
; APPLICANT: Galas, David J.  
; APPLICANT: Kovacevich, Brian  
; APPLICANT: Mulligan, John T.  
; APPLICANT: Paepfer, Bryan W.  
; APPLICANT: Van Ness, Jeffrey  
; APPLICANT: Winkler, David G.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING  
; TITLE OF INVENTION: BONE MINERALIZATION  
; FILE REFERENCE: 240083.508  
; CURRENT APPLICATION NUMBER: US/09/449,218D  
; CURRENT FILING DATE: 1999-11-24  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 43  
; LENGTH: 267  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-449-218D-43

Query Match 52.9%; Score 37; DB 4; Length 267;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 THRIHES 9  
: : : : :  
DB 154 SHEVHMET 161

RESULT 10  
US-09-668-529A-43  
; Sequence 43, Application US/09668529A  
; Patent No. 6489445  
; GENERAL INFORMATION:  
; APPLICANT: Brunkow, Mary E.  
; APPLICANT: Galas, David J.  
; APPLICANT: Kovacevich, Brian  
; APPLICANT: Mulligan, John T.  
; APPLICANT: Paepfer, Bryan W.  
; APPLICANT: Van Ness, Jeffrey  
; APPLICANT: Winkler, David G.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING BONE  
; TITLE OF INVENTION: MINERALIZATION  
; FILE REFERENCE: 240083.508D1  
; CURRENT APPLICATION NUMBER: US/09/668,529A  
; CURRENT FILING DATE: 2000-09-21  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 43  
; LENGTH: 267  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-668-529A-43

Query Match 52.9%; Score 37; DB 4; Length 267;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 THRIHES 9  
: : : : :  
DB 154 SHEVHMET 161

RESULT 11

US-09-668-037A-43  
; Sequence 43, Application US/09668037A  
; Patent No. 6495736  
; GENERAL INFORMATION:  
; APPLICANT: Brunkow, Mary E.  
; APPLICANT: Galas, David J.  
; APPLICANT: Kovacevich, Brian  
; APPLICANT: Mulligan, John T.  
; APPLICANT: Paepfer, Bryan W.  
; APPLICANT: Van Ness, Jeffrey  
; APPLICANT: Winkler, David G.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING BONE  
; TITLE OF INVENTION: MINERALIZATION  
; FILE REFERENCE: 240083.508D4  
; CURRENT APPLICATION NUMBER: US/09/668,037A  
; CURRENT FILING DATE: 2000-09-21  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 43  
; LENGTH: 267  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-668-037A-43

Query Match 52.9%; Score 37; DB 4; Length 267;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 THRIHES 9  
: : : : :  
DB 154 SHEVHMET 161

RESULT 12  
US-08-887-997B-2  
; Sequence 2, Application US/08887997B  
; Patent No. 5935852  
; GENERAL INFORMATION:  
; APPLICANT: FOLETTIE, MAXIMILIAN  
; APPLICANT: DEROBERTIS, EDWARD M.  
; TITLE OF INVENTION: Mammalian Cerberus-Like Protein &  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 Cambridgepark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: US  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/887,997B  
; FILING DATE: 03-JUL--1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LAZAR, STEVEN R  
; REGISTRATION NUMBER: 32,618  
; REFERENCE/DOCKET NUMBER: GI 5290  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 498-8260  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 272 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-887-997B-2

Query Match 52.9%; Score 37; DB 2; Length 272;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 THRIHES 9  
: : : : :  
Db 154 SHEVHER 161

RESULT 13  
US-08-311-731A-291  
; Sequence 291, Application US/08311731A  
; Patent No. 6583266  
; GENERAL INFORMATION:  
; APPLICANT: SMITH, DOUGLAS  
; APPLICANT: MAO, JEN-I  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LABRAE FOR  
; NUMBER OF SEQUENCES: 411  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
; STREET: 600 ATLANTIC AVENUE  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: USA  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311.731A  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GATES, EDWARD R.  
; REGISTRATION NUMBER: 31,616  
; REFERENCE/DOCKET NUMBER: C0044/7125  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/720-2441  
; TELEFAX: 617/720-3500  
; INFORMATION FOR SEQ ID NO: 291:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 526 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Mycobacterium leprae  
US-08-311-731A-291  
Query Match 52.9%; Score 37; DB 4; Length 526;  
Best Local Similarity 50.0%; Pred. No. 2.2e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 HRIHESASL 12  
: : : : :  
Db 165 HLVMWATANL 174

RESULT 14  
US-08-311-731A-71  
; Sequence 71, Application US/08311731A  
; Patent No. 6583266  
; GENERAL INFORMATION:  
; APPLICANT: SMITH, DOUGLAS  
; APPLICANT: MAO, JEN-I  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LABRAE FOR

; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 411  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
; STREET: 600 ATLANTIC AVENUE  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: USA  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311.731A  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GATES, EDWARD R.  
; REGISTRATION NUMBER: 31,616  
; REFERENCE/DOCKET NUMBER: C0044/7125  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/720-2441  
; TELEFAX: 617/720-3500  
; INFORMATION FOR SEQ ID NO: 71:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 538 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: MYCOBACTERIUM LEPRAE  
US-08-311-731A-71  
Query Match 52.9%; Score 37; DB 4; Length 538;  
Best Local Similarity 50.0%; Pred. No. 2.2e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 HRIHESASL 12  
: : : : :  
Db 177 HLVMWATANL 186

RESULT 15  
US-09-413-814-68  
; Sequence 68, Application US/09413814  
; Patent No. 6225064  
; GENERAL INFORMATION:  
; APPLICANT: Geselshaft fuer Biotechnologische Forschung mbH  
; APPLICANT: Bristol-Myers Squibb, Co.  
; APPLICANT: Beyer, Stefan  
; APPLICANT: Bloecker, Helmut  
; APPLICANT: Brandt, Petra  
; APPLICANT: Cino, Paul M  
; APPLICANT: Dougherty, Brian A  
; APPLICANT: Goldberg, Steven L  
; APPLICANT: Hofle, Gerhard  
; APPLICANT: Mueller, Joachim  
; APPLICANT: Reichenbach, Hans  
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or  
; TITLE OF INVENTION: heteropolypeptide compounds  
; FILE REFERENCES: PCT/US 99/23535  
; CURRENT APPLICATION NUMBER: US/09/413.814  
; EARLIER FILING DATE: 1999-10-07  
; EARLIER APPLICATION NUMBER: DE 198 46 493.2  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 68  
; LENGTH: 700  
; TYPE: PRT

ORGANISM: Sorangium cellulosum  
US-09-413-814-68

Query Match 52.9%; Score 37; DB 3; Length 700;  
Best Local Similarity 60.0%; Pred. No. 2.9e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 THRIHESASL 13  
:|:|:|:|  
DB 269 RLHMDMAQL 278

RESULT 16  
US-08-839-710-3  
Sequence 3, Application US/08839710  
Patent No. 5776698  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Goli, Surya K.  
APPLICANT: Streeter, David G.  
TITLE OF INVENTION: NEW REGULATOR OF GENE  
TITLE OF INVENTION: TRANSCRIPTION  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/839,710  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0220 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 93 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 202344  
US-08-839-710-3

Query Match 51.4%; Score 36; DB 1; Length 93;  
Best Local Similarity 45.5%; Pred. No. 55;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 THRIHESASL 12  
:|:|:|:|  
DB 20 SHRVTEGAEV 30

RESULT 17  
US-09-066-262-3  
Sequence 3, Application US/09066262  
Patent No. 5965706  
GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.  
APPLICANT: Goli, Surya K.  
APPLICANT: Streeter, David G.  
TITLE OF INVENTION: NEW REGULATOR OF GENE  
TITLE OF INVENTION: TRANSCRIPTION  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/066,262  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/839,710  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0220 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 93 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 202344  
US-09-066-262-3

Query Match 51.4%; Score 36; DB 2; Length 93;  
Best Local Similarity 45.5%; Pred. No. 55;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 THRIHESASL 12  
:|:|:|:|  
DB 20 SHRVTEGAEV 30

RESULT 18  
US-08-959-004-10  
Sequence 10, Application US/08959004  
Patent No. 6197543  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Vye, Henry  
APPLICANT: Corley, Neil C.  
APPLICANT: Lal, Preeti  
APPLICANT: Shah, Purvi  
APPLICANT: Kaser, Matthew  
TITLE OF INVENTION: HUMAN VESICLE MEMBRANE PROTEIN-LIKE  
TITLE OF INVENTION: PROTEINS  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304



COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/959.004  
FILING DATE: Herewith  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0414 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 625 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1665777  
US-08-959-004-10

Query Match 51.4%; Score 36; DB 3; Length 625;  
Best Local Similarity 36.4%; Pred. No. 3.7e+02;  
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 THRIHESASL 12  
: : : : :  
Db 234 TYSVHWESDI 244

RESULT 19  
US-07-646-537B-2  
Sequence 2, Application US/07646537B  
Patent No. 534864  
GENERAL INFORMATION:  
APPLICANT: Bairdaci, Mariano  
TITLE OF INVENTION: Vav Proto-Oncogene Protein  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bristol-Myers Squibb Company  
STREET: P.O. Box 4000  
CITY: Princeton  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 08543-4000  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/646.537B  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gaul, Timothy J.  
REGISTRATION NUMBER: 33,111  
REFERENCE/DOCKET NUMBER: DC10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (609) 921-5901  
TELEFAX: (609) 921-4526  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:

LENGTH: 844 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-646-537B-2

Query Match 51.4%; Score 36; DB 1; Length 844;  
Best Local Similarity 45.5%; Pred. No. 5e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 THRIHESASL 12  
: : : : :  
Db 20 SHRVTEGAQV 30

RESULT 20  
US-09-214-631-7  
Sequence 7, Application US/09214631  
Patent No. 6413730  
GENERAL INFORMATION:  
APPLICANT: Holland, Sacha  
APPLICANT: Mamailu, Geraldine  
APPLICANT: Pawson, Tony  
TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED  
BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR  
TITLE OF INVENTION: TYROSINE KINASES  
FILE REFERENCE: 11757.23USWO  
CURRENT APPLICATION NUMBER: US/09/214.631  
CURRENT FILING DATE: 1999-03-12  
EARLIER APPLICATION NUMBER: PCT/CA97/00473  
EARLIER FILING DATE: 1997-07-04  
EARLIER APPLICATION NUMBER: 60/021.272  
EARLIER FILING DATE: 1996-07-05  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7  
LENGTH: 233  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-214-631-7

Query Match 50.0%; Score 35; DB 4; Length 233;  
Best Local Similarity 57.1%; Pred. No. 2e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 THRIHWE 8  
: : : : :  
Db 146 THNLHWK 152

RESULT 21  
US-08-299-567-5  
Sequence 5, Application US/08299567  
Patent No. 574703  
GENERAL INFORMATION:  
APPLICANT: Davis, et al.  
TITLE OF INVENTION: METHOD OF ENHANCING THE BIOLOGICAL  
ACTIVITY OF EPH FAMILY LIGANDS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
STREET: 777 Old Saw Mill River Road  
CITY: Tarrytown  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10591-6707  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/299.567

FILING DATE: 01-SEP-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kempier, Gail M.  
REGISTRATION NUMBER: 32,143  
REFERENCE/DOCKET NUMBER: REG 290  
TELEPHONE: 914-345-7400  
TELEFAX: 914-345-7721  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 234 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-299-567-5

Query Match 50.0%; Score 35; DB 1; Length 234;  
Best Local Similarity 57.1%; Pred. No. 2e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 THRIHWE 8  
DB 147 THNLRHWK 153

RESULT 22  
US-08-240-124-2  
Sequence 2, Application US/08240124  
Patent No. 5516658  
GENERAL INFORMATION:  
APPLICANT: BECKMANN, M. P.  
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE  
NUMBER OF SEQUENCES: 4  
TITLE OF INVENTION: RECEPTOR HEK  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMUNEX CORPORATION  
STREET: 51 UNIVERSITY STREET  
CITY: SEATTLE  
STATE: WASHINGTON  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: Apple Macintosh  
SOFTWARE: Microsoft Word for Apple, Version 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/240,124  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/161,132  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/114,426  
FILING DATE: 30-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/109,745  
FILING DATE: 20-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: SESE, KATHRYN A.  
REGISTRATION NUMBER: 32,172  
REFERENCE/DOCKET NUMBER: 2814-C  
TELEPHONE: (206) 587-0430  
TELEFAX: (206) 233-0644  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 238 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-240-124-2

Query Match 50.0%; Score 35; DB 1; Length 238;  
Best Local Similarity 57.1%; Pred. No. 2e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 THRIHWE 8  
DB 151 THNLRHWK 157

RESULT 23  
US-08-453-943-2  
Sequence 2, Application US/08453943  
Patent No. 5738844  
GENERAL INFORMATION:  
APPLICANT: BECKMANN, M. P.  
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE  
NUMBER OF SEQUENCES: 4  
TITLE OF INVENTION: RECEPTOR HEK  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMUNEX CORPORATION  
STREET: 51 UNIVERSITY STREET  
CITY: SEATTLE  
STATE: WASHINGTON  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: Apple Macintosh  
SOFTWARE: Microsoft Word for Apple, Version 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,943  
FILING DATE: 30-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/240,124  
FILING DATE: 09-MAY-1994  
APPLICATION NUMBER: US 08/161,132  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/114,426  
FILING DATE: 30-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/109,745  
FILING DATE: 20-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: SESE, KATHRYN A.  
REGISTRATION NUMBER: 32,172  
REFERENCE/DOCKET NUMBER: 2814-C  
TELEPHONE: (206) 587-0430  
TELEFAX: (206) 233-0644  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 238 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-453-943-2

Query Match 50.0%; Score 35; DB 1; Length 238;  
Best Local Similarity 57.1%; Pred. No. 2e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 THRIHWE 8



Query Match 50.0%; Score 35; DB 4; Length 332;  
Best Local Similarity 53.8%; Pred. No. 2.8e+02;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ITHRIHWESASL 13  
||| : ||  
||| : ||  
DB 266 ITHPIEMQLKYL 278

RESULT 27  
US-09-252-991A-28288  
; Sequence 28288, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 28288  
; LENGTH: 356  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-28288

Query Match 50.0%; Score 35; DB 4; Length 356;  
Best Local Similarity 85.7%; Pred. No. 3.1e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 HWESASL 12  
||| : ||  
||| : ||  
DB 19 HWSASL 25

RESULT 28  
US-09-252-991A-31968  
; Sequence 31968, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 31968  
; LENGTH: 542  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-31968

Query Match 50.0%; Score 35; DB 4; Length 542;  
Best Local Similarity 40.0%; Pred. No. 4.7e+02;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 HRIHWESASL 12  
||| : ||  
||| : ||  
DB 236 HRLHWAAGV 245

RESULT 29

US-08-501-572-1  
; Sequence 1, Application US/08501572  
; Patent No. 6063623  
; GENERAL INFORMATION:

; APPLICANT: Koepsell, Hermann  
; APPLICANT: Grundeman, Dirk  
; TITLE OF INVENTION: Transport of protein Which Effects The  
; TITLE OF INVENTION: Transport Of Cationic Xenobiotics and/or Pharmaceuticals,  
; TITLE OF INVENTION: DNA Sequences Encoding It And Their Use.  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner  
; STREET: 1300 I Street, N.W., Suite 700  
; CITY: Washington  
; STATE: D.C.

; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/501,572  
; FILING DATE:

; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Toohey, Kimberlin M  
; REGISTRATION NUMBER: 35,391  
; REFERENCE/DOCKET NUMBER: 02481.1453-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)408-4400  
; TELEPHONE: (202)408-4000  
; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 556 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-501-572-1

Query Match 50.0%; Score 35; DB 3; Length 556;  
Best Local Similarity 41.7%; Pred. No. 4.8e+02;  
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 ITHRIHWESASL 12  
||| : ||  
||| : ||  
DB 424 IPIHEHMLNVTL 435

RESULT 30  
US-09-040-444-1  
; Sequence 1, Application US/09040444  
; Patent No. 6063766  
; GENERAL INFORMATION:

; APPLICANT: Koepsell, Hermann  
; APPLICANT: Grundeman, Dirk  
; TITLE OF INVENTION: Transport protein Which Effects The  
; TITLE OF INVENTION: Transport Of Cationic Xenobiotics and/or Pharmaceuticals,  
; TITLE OF INVENTION: DNA Sequences Encoding It And Their Use.  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.  
; STREET: 1300 I Street, N.W., Suite 700  
; CITY: Washington  
; STATE: D.C.

; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/040,444  
FILING DATE: March 18, 1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: O'Connor, Steven P  
REGISTRATION NUMBER: 41,225  
REFERENCE/DOCKET NUMBER: 2481.1453-01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 408-4000  
TELEFAX: (202) 408-4400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 556 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-040-444-1

Query Match 50.0%; Score 35; DB 3; Length 556;  
Best Local Similarity 41.7%; Pred. No. 4.8e+02;  
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 ITRHWESAST 12  
| | | | | | | | | | | | | | | | | | | | | |  
DB 424 IPHELMWLVTL 435

RESULT 31  
US-09-491-522-7  
Sequence 7, Application US/09491522  
Patent No. 6428998  
GENERAL INFORMATION:  
APPLICANT: Colige, Alain  
APPLICANT: Lapierre, Charles M.  
APPLICANT: Prockop, Darwin J.  
TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,  
TITLE OF INVENTION: AND THE PRODUCTION, METHODS AND USES THEREOF  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds, LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2811  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/491,522  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/886,333  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Abrams, Samuel B  
REGISTRATION NUMBER: 30,605  
REFERENCE/DOCKET NUMBER: 8389-0060-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-493-4935  
TELEFAX: 650-493-5556  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 566 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-491-522-7

Query Match 50.0%; Score 35; DB 4; Length 566;  
Best Local Similarity 55.6%; Pred. No. 4.9e+02;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 HRHWESAS 11  
| | | | | | | | | | | | | | | | | | | | | |  
DB 444 HRFHWSRCS 452

RESULT 32  
US-09-252-991A-18995  
Sequence 18995, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 18995  
LENGTH: 595  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18995

Query Match 50.0%; Score 35; DB 4; Length 595;  
Best Local Similarity 80.0%; Pred. No. 5.1e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 HRHW 7  
| | | | | | | | | | | | | | | | | | | | | |  
DB 364 HRVHW 368

RESULT 33  
US-09-107-532A-7065  
Sequence 7065, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
TITLE OF INVENTION: ENTEROCOCCUS FACIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571

```

; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Denekle
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-8277
; TELEFAX: (781)893-5007
; INFORMATION FOR SEQ ID NO: 7065:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 796 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...796
; SEQUENCE DESCRIPTION: SEQ ID NO: 7065:
US-09-107-532A-7065

Query Match
Best Local Similarity 50.0%; Score 35; DB 4; Length 796;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 THRIWES 9
DB 113 TVRFHWES 120

RESULT 34
US-09-252-991A-26112
; Sequence 26112, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfeld et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26112
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26112

Query Match
Best Local Similarity 50.0%; Score 35; DB 4; Length 860;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 THRIWESASL 12
DB 93 LTHRLSPKQASL 104

RESULT 35
US-09-491-522-11
; Sequence 11, Application US/09491522
; Patent No. 6428998
; GENERAL INFORMATION:
; APPLICANT: Colige, Alain
; APPLICANT: Prockop, Darwin J.
; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
; AND THE PRODUCTION, METHODS AND USES THEREOF
```

```

; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/491,522
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/886,333
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Abrams, Samuel B
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8389-0060-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1205 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-491-522-11

Query Match
Best Local Similarity 50.0%; Score 35; DB 4; Length 1205;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 HRIWESAS 11
DB 438 HRFHWRS 446

RESULT 36
US-09-491-522-5
; Sequence 5, Application US/09491522
; Patent No. 6428998
; GENERAL INFORMATION:
; APPLICANT: Colige, Alain
; APPLICANT: Lapiere, Charles M.
; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
; AND THE PRODUCTION, METHODS AND USES THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/491,522
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
```

APPLICATION NUMBER: 08/886,333  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Abrams, Samuel B  
REGISTRATION NUMBER: 30,605  
REFERENCE/DOCKET NUMBER: 8389-0060-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-493-4935  
TELEFAX: 650-493-5556  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1211 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-491-522-5

Query Match 50.0%; Score 35; DB 4; Length 1211;  
Best Local Similarity 55.6%; Pred. No. 1e+03;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 HRHWSAS 11  
Db 444 HRFHWSRCS 452

RESULT 37  
US-08-889-013C-6  
Sequence 6, Application US/08889013C  
Patent No. 5919640  
GENERAL INFORMATION:  
APPLICANT: TIKKANEN, KARIINA  
TITLE OF INVENTION: STREPTOCOCCUS SUIS ADHESIN PROTEIN AND  
TITLE OF INVENTION: METHOD FOR PRODUCING IT  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHAYE P. C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/889,013C  
FILING DATE: 07-JUL-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FI 930413  
FILING DATE: 29-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAWFORD, ARTHUR R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 227-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 105 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus suis

US-08-889-013C-6

Query Match 48.6%; Score 34; DB 2; Length 105;  
Best Local Similarity 57.1%; Pred. No. 1.3e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ITHRIHW 7  
Db 19 ILHQVHW 25

RESULT 38  
US-09-107-532A-3813  
Sequence 3813, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
INFORMATION FOR SEQ ID NO: 3813:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 126 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...126  
SEQUENCE DESCRIPTION: SEQ ID NO: 3813:  
US-09-107-532A-3813

Query Match 48.6%; Score 34; DB 4; Length 126;  
Best Local Similarity 62.5%; Pred. No. 1.5e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 HWSASLI 13  
Db 93 HWEITSVL 100

RESULT 39  
US-08-729-103-4

; Sequence 4, Application US/08729103  
; Patent No. 5837841  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Goll, Surya K.  
; TITLE OF INVENTION: NOVEL HUMAN REG PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/729,103  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-845-4166  
; TELEFAX: 415-845-0555  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 166 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 474306  
; US-08-729-103-4

Query Match 48.6%; Score 34; DB 2; Length 166;  
Best Local Similarity 60.0%; Pred. No. 2e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 RHMESASLL 13  
| | | | |  
Db 110 RWHMSSGLV 119

RESULT 40  
US-09-252-991A-19372  
; Sequence 19372, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 19372  
; LENGTH: 226  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-19372  
Query Match 48.6%; Score 34; DB 4; Length 226;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 HRIHW 7  
| | | | |  
Db 18 HRLHW 22

Search completed: August 28, 2003, 14:14:56  
Job time : 30 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 28, 2003, 14:12:57 ; Search time 57 Seconds  
(without alignments)  
31.196 Million cell updates/sec

Title: US-09-845-738C-1  
Perfect score: 70  
Sequence: 1 ITHRIHMSASL 13

Scoring table:  
BLAST62  
Gapop 10.0, Gapext 0.5

Searched: 510680 seqs, 136781880 residues  
Total number of hits satisfying chosen parameters: 510680

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PTC\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PTCUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09D\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	70	100.0	14	US-09-845-730-1	Sequence 1, Appli
2	70	100.0	16	US-09-846-345-1	Sequence 1, Appli
3	70	100.0	17	US-09-846-346-1	Sequence 1, Appli
4	70	100.0	1663	US-09-875-519A-22	Sequence 22, Appli
5	70	100.0	1663	US-09-842-758-41	Sequence 41, Appli
6	66	94.3	12	US-09-846-349-1	Sequence 1, Appli
7	61	87.1	11	US-09-845-715-1	Sequence 1, Appli
8	53	75.7	10	US-09-845-731-1	Sequence 1, Appli
9	52	74.3	1661	US-09-842-758-42	Sequence 42, Appli
10	39	55.7	321	US-10-183-116-45	Sequence 45, Appli
11	39	55.7	322	US-10-183-116-47	Sequence 47, Appli
12	39	55.7	387	US-10-156-761-12347	Sequence 12347, A
13	38	54.3	55	US-10-074-475-209	Sequence 209, App
14	38	54.3	268	US-10-183-116-79	Sequence 79, Appli
15	38	54.3	530	US-10-205-219-121	Sequence 121, App

16	38	54.3	2012	10	US-09-808-602-68	Sequence 68, Appli
17	38	54.3	2012	11	US-09-800-198-57	Sequence 57, Appli
18	37	52.9	40	9	US-09-864-761-35988	Sequence 35988, A
19	37	52.9	69	9	US-09-864-761-44965	Sequence 44965, A
20	37	52.9	103	9	US-09-864-761-34487	Sequence 34487, A
21	37	52.9	134	9	US-09-864-761-46114	Sequence 46114, A
22	37	52.9	151	12	US-10-002-631C-135	Sequence 135, App
23	37	52.9	267	10	US-09-089-818B-8	Sequence 8, Appli
24	37	52.9	267	12	US-10-139-814-12	Sequence 12, Appli
25	37	52.9	267	12	US-10-286-152A-52	Sequence 52, Appli
26	37	52.9	267	14	US-10-044-716-12	Sequence 12, Appli
27	37	52.9	272	10	US-09-887-552A-2	Sequence 2, Appli
28	37	52.9	272	10	US-09-089-818B-2	Sequence 2, Appli
29	37	52.9	292	15	US-10-106-698-4945	Sequence 4945, Ap
30	37	52.9	301	15	US-10-097-340-93	Sequence 93, Appli
31	37	52.9	385	14	US-10-139-262-2	Sequence 2, Appli
32	37	52.9	385	15	US-10-255-969-2	Sequence 2, Appli
33	37	52.9	390	11	US-09-764-891-4103	Sequence 4103, Ap
34	37	52.9	390	15	US-10-103-313-418	Sequence 418, App
35	37	52.9	390	15	US-10-103-313-541	Sequence 541, App
36	37	52.9	433	14	US-10-139-262-6	Sequence 6, Appli
37	37	52.9	433	15	US-10-255-969-6	Sequence 6, Appli
38	37	52.9	440	14	US-10-139-262-4	Sequence 4, Appli
39	37	52.9	440	15	US-10-255-969-4	Sequence 4, Appli
40	37	52.9	451	10	US-09-938-330-2	Sequence 2, Appli
41	37	52.9	486	10	US-09-938-330-6	Sequence 6, Appli
42	37	52.9	486	15	US-10-270-333-75	Sequence 75, Appli
43	37	52.9	545	11	US-09-374-046A-26	Sequence 26, Appli
44	37	52.9	609	9	US-09-815-242-11758	Sequence 11758, A
45	37	52.9	1156	14	US-10-014-070-5	Sequence 5, Appli

#### ALIGNMENTS

RESULT 1  
US-09-845-730-1  
Sequence 1, Application US/09845730  
Patent No. US20020169278A1  
GENERAL INFORMATION:  
APPLICANT: Jackowski, George  
TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR  
FILE REFERENCE: 2132.042  
CURRENT APPLICATION NUMBER: US/09/845.730  
CURRENT FILING DATE: 2001-04-30  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-845-730-1

Query Match 100.0%; Score 70; DB 10; Length 14;  
Best Local Similarity 100.0%; Pred. NO. 7.3e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITHRIHMSASL 13  
|||||||  
Db 2 ITHRIHMSASL 14

RESULT 2  
US-09-846-345-1  
Sequence 1, Application US/09846345  
Patent No. US20020161182A1  
GENERAL INFORMATION:  
APPLICANT: Jackowski, George  
TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR  
FILE REFERENCE: 2132.045  
CURRENT APPLICATION NUMBER: US/09/846.345

/ CURRENT FILING DATE: 2001-04-30  
/ NUMBER OF SEQ ID NOS: 1  
/ SOFTWARE: PatentIn version 3.1  
/ SEQ ID NO: 1  
/ LENGTH: 16  
/ TYPE: PRT  
/ ORGANISM: Homo sapiens  
US-09-846-345-1

Query Match 100.0%; Score 70; DB 10; Length 16;  
Best Local Similarity 100.0%; Pred. No. 8.3e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITHRHIMESASLL 13  
|||||  
Db 4 ITHRHIMESASLL 16

RESULT 3  
US-09-846-346-1  
/ Sequence 1, Application US/09846346  
/ Patent No. US20020160532A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Jackowski, George  
/ TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR  
/ FILE REFERENCE: OF 1998 DALTONS  
/ CURRENT APPLICATION NUMBER: US/09/846,346  
/ CURRENT FILING DATE: 2001-04-30  
/ NUMBER OF SEQ ID NOS: 1  
/ SOFTWARE: PatentIn version 3.1  
/ SEQ ID NO: 1  
/ LENGTH: 17  
/ TYPE: PRT  
/ ORGANISM: Homo sapiens  
US-09-846-346-1

Query Match 100.0%; Score 70; DB 10; Length 17;  
Best Local Similarity 100.0%; Pred. No. 8.8e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITHRHIMESASLL 13  
|||||  
Db 4 ITHRHIMESASLL 16

RESULT 4  
US-09-875-519A-22  
/ Sequence 22, Application US/09875519A  
/ Patent No. US20020068059A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Parries, Timothy C.  
/ APPLICANT: Harrison, Richard A.  
/ TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase  
/ FILE REFERENCE: 4-30443/A/IMU/PCT  
/ CURRENT APPLICATION NUMBER: US/09/875,519A  
/ CURRENT FILING DATE: 2001-06-06  
/ PRIOR APPLICATION NUMBER: PCT/GB97/00603  
/ PRIOR FILING DATE: 1997-03-04  
/ NUMBER OF SEQ ID NOS: 35  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO: 22  
/ LENGTH: 1663  
/ TYPE: PRT  
/ ORGANISM: Homo sapiens  
US-09-875-519A-22

Query Match 100.0%; Score 70; DB 9; Length 1663;  
Best Local Similarity 100.0%; Pred. No. 0.0063;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ITHRHIMESASLL 13  
|||||

Db 1307 ITHRHIMESASLL 1319

RESULT 5  
US-09-842-758-41  
/ Sequence 41, Application US/09842758  
/ Publication No. US20030083244A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Vernet, Corine A. M.  
/ APPLICANT: Fernandes, Elma R.  
/ APPLICANT: Gerlach, Valerie  
/ APPLICANT: Shinkets, Richard A.  
/ APPLICANT: Malyankar, Uziel M.  
/ APPLICANT: Boldog, Ferenc L.  
/ APPLICANT: Zernusen, Bryan D.  
/ APPLICANT: Spyrek, Kimberly A.  
/ APPLICANT: Majumder, Kunud  
/ APPLICANT: Tchernev, Velizar T.  
/ APPLICANT: Padigar, Muraidhara  
/ APPLICANT: Patturajan, Meera  
/ APPLICANT: Burgess, Catherine E.  
/ APPLICANT: Gangolli, Esna A.  
/ APPLICANT: Smithson, Glenda  
/ APPLICANT: Rastelli, Luca  
/ APPLICANT: MacDougall, John R.  
/ APPLICANT: Taupier, Raymond J.  
/ APPLICANT: Grosse, William M.  
/ APPLICANT: Edward, Szekeres S.  
/ APPLICANT: Alsbrook II, John P.  
/ TITLE OF INVENTION: No. US20030083244A1 Proteins and Nucleic Acids Encoding Same  
/ FILE REFERENCE: 15966-783  
/ CURRENT APPLICATION NUMBER: US/09/842,758  
/ CURRENT FILING DATE: 2001-04-25  
/ PRIOR APPLICATION NUMBER: 60/200,158  
/ PRIOR FILING DATE: 2000-04-26  
/ PRIOR APPLICATION NUMBER: 60/200,613  
/ PRIOR FILING DATE: 2000-04-28  
/ PRIOR APPLICATION NUMBER: 60/200,780  
/ PRIOR FILING DATE: 2000-04-28  
/ PRIOR APPLICATION NUMBER: 60/201,006  
/ PRIOR FILING DATE: 2000-05-01  
/ PRIOR APPLICATION NUMBER: 60/201,007  
/ PRIOR FILING DATE: 2000-05-01  
/ PRIOR APPLICATION NUMBER: 60/201,236  
/ PRIOR FILING DATE: 2000-05-01  
/ PRIOR APPLICATION NUMBER: 60/201,238  
/ PRIOR FILING DATE: 2000-05-01  
/ PRIOR APPLICATION NUMBER: 60/201,186  
/ PRIOR FILING DATE: 2000-05-02  
/ PRIOR APPLICATION NUMBER: 60/201,474  
/ PRIOR FILING DATE: 2000-05-03  
/ PRIOR APPLICATION NUMBER: 60/201,508  
/ PRIOR FILING DATE: 2000-05-03  
/ PRIOR APPLICATION NUMBER: 60/220,591  
/ PRIOR FILING DATE: 2000-07-25  
/ PRIOR APPLICATION NUMBER: 60/232,678  
/ PRIOR FILING DATE: 2000-09-15  
/ PRIOR APPLICATION NUMBER: 60/263,217  
/ PRIOR FILING DATE: 2001-01-22  
/ PRIOR APPLICATION NUMBER: 60/265,160  
/ PRIOR FILING DATE: 2001-01-30  
/ NUMBER OF SEQ ID NOS: 113  
/ SOFTWARE: PatentIn Ver. 2.1  
/ SEQ ID NO: 41  
/ LENGTH: 1663  
/ TYPE: PRT  
/ ORGANISM: Homo sapiens  
US-09-842-758-41

Query Match 100.0%; Score 70; DB 11; Length 1663;  
Best Local Similarity 100.0%; Pred. No. 0.0063;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 ITHRIHWESASLL 13
        |||||
Db      1307 ITHRIHWESASLL 1319
```

```

RESULT 6
US-09-846-349-1
: Sequence 1, Application US/09846349
: Patent No. US20020161186A1
: GENERAL INFORMATION:
: APPLICANT: JACOWSKI, George
: TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
: TITLE OF INVENTION: OF 1449 DALTONS
: FILE REFERENCE: 2132.034
: CURRENT APPLICATION NUMBER: US/09/846,349
: CURRENT FILING DATE: 2001-04-30
: NUMBER OF SEQ ID NOS: 1
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO: 1
: LENGTH: 12
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-846-349-1

```

Query Match	94.3%	Score 66	DB 10	Length 12
Best Local Similarity	100.0%	Pred. No. 0.00028		
Matches 12	Conservative 0	Mismatches 0	Indels 0	Gaps 0

```
QY      2 THRIHWESASLL 13
        |||||
Db      1 THRIHWESASLL 12
```

```

RESULT 7
US-09-845-715-1
: Sequence 1, Application US/09845715
: Patent No. US20020161184A1
GENERAL INFORMATION:
APPLICANT: JACKOWSKI, George
: TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
: TITLE OF INVENTION: OF 1348 DALTONS
: FILE REFERENCE: 2132.030
: CURRENT APPLICATION NUMBER: US/09/845,715
: CURRENT FILING DATE: 2001-04-30
: NUMBER OF SEQ ID NOS: 1
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 11
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-845-715-1

```

Query Match	87.1%	Score 61	DB 10	Length 11
Best Local Similarity	100.0%	Pred. No. 0.0016		
Matches 11: Conservative	0	Mismatches	0	Indels 0
				Gaps 0

```
Qy      3 HRIHWESASLL 13
         |||||
Db      1 HRIHWESASLL 11
```

```

RESULT 8
US-09-845-731-1
/ Sequence 1, Application US/09845731
/ Publication No. US20030004307A1
/ GENERAL INFORMATION:
/ APPLICANT: Jackowski, George
/ TITLE OF INVENTION: Biopolymer Marker Indicative Of Disease State Having A Molecular
/ FILE OF INVENTION: Of 1211 Daltons
/ FILE REFERENCE: 2132,029
/ CURRENT APPLICATION NUMBER: US/09/845,731
/ CURRENT FILING DATE: 2001-04-30
/ NUMBER OF SEQ ID NOS: 1

```

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 10
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-845-731-1

```

Query Match	75.74;	Score 53;	DB 11;	Length 10;
Best Local Similarity	100.04;	Pred. No. 0.028;		
Matches 10;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      4 RHWESASLL 13
        |||||
Db      1 RHWESASLL 10
```

```

RESULT 9
US-09-842-758-42
: Sequence 42, Application US/09842758
: Publication No. US20030083244A1
: GENERAL INFORMATION:
: APPLICANT: Vernet, Corine A. M.
: APPLICANT: Fernandes, Elma R
: APPLICANT: Gerlach, Valerie
: APPLICANT: Shimkets, Richard A
: APPLICANT: Malyankar, Uriel M
: APPLICANT: Boldog, Ferenc L
: APPLICANT: Zernhusen, Bryan D
: APPLICANT: Spytek, Kimberly A
: APPLICANT: Majumder, Kumud
: APPLICANT: Tchernev, Velizar T
: APPLICANT: Padigaru, Muralidhara
: APPLICANT: Patuturajan, Meera
: APPLICANT: Burgess, Catherine E
: APPLICANT: Gangolli, Esha A
: APPLICANT: Smithson, Glenda
: APPLICANT: Rastelli, Luca
: APPLICANT: Macdougall, John R
: APPLICANT: Tauplier, Raymond J
: APPLICANT: Grosee, William M
: APPLICANT: Edward, Szekeres S
: APPLICANT: Alsobrook II, John P
: TITLE OF INVENTION: No. US20030083244A1 Proteins and Nucleic Acids Encoding Same
: FILE REFERENCE: 15966-783
: CURRENT APPLICATION NUMBER: US/09/842,758
: CURRENT FILING DATE: 2001-04-25
: PRIOR APPLICATION NUMBER: 60/200,158
: PRIOR FILING DATE: 2000-04-26
: PRIOR APPLICATION NUMBER: 60/200,613
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,780
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/201,006
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/201,007
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/201,236
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/201,238
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/201,186
: PRIOR FILING DATE: 2000-05-02
: PRIOR APPLICATION NUMBER: 60/201,474
: PRIOR FILING DATE: 2000-05-03
: PRIOR APPLICATION NUMBER: 60/201,508
: PRIOR FILING DATE: 2000-05-03
: PRIOR APPLICATION NUMBER: 60/220,591
: PRIOR FILING DATE: 2000-07-25
: PRIOR APPLICATION NUMBER: 60/232,678
: PRIOR FILING DATE: 2000-09-15
: PRIOR APPLICATION NUMBER: 60/263,217
: PRIOR FILING DATE: 2001-01-22
: PRIOR APPLICATION NUMBER: 60/265,160

```

```

; GENERAL INFORMATION:
; APPLICANT: Salceda, Susanaana
; APPLICANT: Macina, Robertt
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Hervee
; APPLICANT: Karra, Kalpanana

```

APPLICANT: Cafferkey, Robert  
APPLICANT: Sun, Yongsung  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific  
FILE REFERENCE: DEX-0313  
CURRENT APPLICATION NUMBER: US/10/074,475  
CURRENT FILING DATE: 2002-02-13  
PRIOR APPLICATION NUMBER: 60/266,292  
PRIOR FILING DATE: 2001-02-13  
NUMBER OF SEQ ID NOS: 295  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 209  
LENGTH: 55  
TYPE: PRT  
ORGANISM: Homo sapien  
US-10-074-475-209

Query Match 54.3%; Score 38; DB 15; Length 55;  
Best Local Similarity 57.1%; Pred. No. 34;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHW 7  
Db 37 VTHKQHW 43

RESULT 14  
US-10-183-116-79  
Sequence 79, Application US/10183116  
Publication No. US20030092035A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, David J.  
APPLICANT: Dong, Xinzhang  
APPLICANT: Zylka, Mark  
APPLICANT: Simon, Melvin  
APPLICANT: Han, Sang-kyou  
TITLE OF INVENTION: PAIN SIGNALING MOLECULES  
FILE REFERENCE: CALTE 4C1CP1  
CURRENT APPLICATION NUMBER: US/10/183,116  
CURRENT FILING DATE: 2002-06-26  
PRIOR APPLICATION NUMBER: US 60/222,344  
PRIOR FILING DATE: 2000-08-01  
PRIOR APPLICATION NUMBER: US 60/202,027  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: US 09/704,707  
PRIOR FILING DATE: 2000-11-03  
PRIOR APPLICATION NUMBER: US 60/285,493  
PRIOR FILING DATE: 2001-04-19  
PRIOR APPLICATION NUMBER: US 09/849,869  
PRIOR FILING DATE: 2001-05-04  
NUMBER OF SEQ ID NOS: 109  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 79  
LENGTH: 268  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-183-116-79

Query Match 54.3%; Score 38; DB 15; Length 268;  
Best Local Similarity 54.5%; Pred. No. 1.5e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 HRIHESASL 13  
Db 250 HRLKQSLKTL 260

RESULT 15  
US-10-205-219-121  
Sequence 121, Application US/10205219  
Publication No. US20030138603A1  
GENERAL INFORMATION:

APPLICANT: Warner-Lambert Company  
APPLICANT: Lee, Kevin  
APPLICANT: Dixon, Aletair  
APPLICANT: Brookspank, Robert  
APPLICANT: Pinnock, Robert  
TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain  
FILE REFERENCE: WL-A-018200  
CURRENT APPLICATION NUMBER: US/10/205,219  
CURRENT FILING DATE: 2002-07-24  
PRIOR APPLICATION NUMBER: GB 0118354.0  
PRIOR FILING DATE: 2001-07-27  
NUMBER OF SEQ ID NOS: 197  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 121  
LENGTH: 530  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: EP70-P-180  
US-10-205-219-121

Query Match 54.3%; Score 38; DB 12; Length 530;  
Best Local Similarity 75.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 THRIHWS 9  
Db 216 SHRIHWS 223

RESULT 16  
US-09-808-602-68  
Sequence 68, Application US/09808602  
Patent No. US2002015115A1  
GENERAL INFORMATION:  
APPLICANT: Vernet, Corine A  
APPLICANT: Fernandes, Elma  
APPLICANT: Shimkets, Richard A  
APPLICANT: Heitman, John L  
APPLICANT: Majumder, Kunud  
APPLICANT: Mishra, Vishnu  
APPLICANT: Mezes, Peter S  
APPLICANT: MacDougall, John  
TITLE OF INVENTION: No. US2002015115A1e1 Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 15966-697 CIP  
CURRENT APPLICATION NUMBER: US/09/808,602  
CURRENT FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 09/800,198  
PRIOR FILING DATE: 2001-03-05  
PRIOR APPLICATION NUMBER: 60/186,596  
PRIOR FILING DATE: 2000-03-03  
NUMBER OF SEQ ID NOS: 114  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 68  
LENGTH: 2012  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-808-602-68

Query Match 54.3%; Score 38; DB 10; Length 2012;  
Best Local Similarity 45.5%; Pred. No. 9.8e+02;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 ITHRIHWSAS 11  
Db 1702 VTHTVHYOSVS 1712

RESULT 17  
US-09-800-198-57  
Sequence 57, Application US/09800198  
Publication No. US20030087816A1  
GENERAL INFORMATION:

```

; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkens, Richard A
; APPLICANT: Hermann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 2012
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-800-198-57

Query Match          54.3%; Score 38; DB 11; Length 2012;
Best Local Similarity 45.5%; Pred. No. 9,8e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      1 ITRHWMESAS 11
Db      1702 VTHVHVSVS 1712

RESULT 18
US-09-864-761-35988
; Sequence 35988. Application US/09864761
; Patent No. US2002048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666.
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35988
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
;
; OTHER INFORMATION: MAP TO AL079338.9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.99
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EST_HUMAN HIT: BF541030.1, EVALUO 7.00e-17
; OTHER INFORMATION: SWISSPROT HIT: P14336, EVALUO 3.60e+00
;
US-09-864-761-35988

Query Match          52.9%; Score 37; DB 9; Length 40;
Best Local Similarity 45.5%; Pred. No. 37;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      3 HRHWMESASL 13
Db      7 HLLHWMESKVI 17

RESULT 19
US-09-864-761-44965
; Sequence 44965. Application US/09864761
; Patent No. US2002048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44965
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC016498.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.83
; OTHER INFORMATION: EST_HUMAN HIT: BE877915.1, EVALUATE 1.10e-02
US-09-864-761-44965
```

```
Query Match          52.9%; Score 37; DB 9; Length 69;
Best Local Similarity 75.0%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      6 HWESASLL 13
        ||| |||
Db      25 HWQGSALL 32
```

```
RESULT 20
US-09-864-761-34487
; Sequence 34487, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34487
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006548.19
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: SWISSPROT HIT: P41064, EVALUATE 1.30e+00
; OTHER INFORMATION: EST_HUMAN HIT: AU140898.1, EVALUATE 9.00e-26
US-09-864-761-34487
```

```
Query Match          52.9%; Score 37; DB 9; Length 103;
Best Local Similarity 70.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Oy      2 THRHWESAS 11
        ||| |||
Db      58 THRHWMSLAS 67
```

```
RESULT 21
US-09-864-761-46114
; Sequence 46114, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46114
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006548.20
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8
; OTHER INFORMATION: EST HUMAN HIT: AUI40898.1, EVALUE 6.00e-37
; OTHER INFORMATION: SWISSPROT HIT: Q09312, EVALUE 2.00e+00
; US-09-864-761-46114

Query Match          52.9%; Score 37; DB 9; Length 134;
Best Local Similarity 70.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy      2 THRIHWSAS 11
        |||||
Db      63 THRIHWSLAS 72
```

```
RESULT 22
US-10-002-631C-135
; Sequence 135, Application US/10002631C
; Publication No. US20030157486A1
; GENERAL INFORMATION:
; APPLICANT: Muenter, Matthew M.
; TITLE OF INVENTION: METHODS TO IDENTIFY SIGNAL SEQUENCES
; FILE REFERENCE: A34943 090495.0243
; CURRENT APPLICATION NUMBER: US/10/002,631C
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/300,309
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 135
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (136)...(136)
```

```
; OTHER INFORMATION: Xaa = any amino acid
US-10-002-631C-135
```

```
Query Match          52.9%; Score 37; DB 12; Length 151;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      3 THRIHWS 9
        |||||
Db      47 THRIHWS 53
```

```
RESULT 23
US-09-089-818B-8
; Sequence 8, Application US/09089818B
; Patent No. US20020164682A1
; GENERAL INFORMATION:
; APPLICANT: Folletle, Maximillian
; APPLICANT: Derobertis, Edward M.
; TITLE OF INVENTION: Mammalian Cerberus-Like Protein &
; TITLE OF INVENTION: Compositions
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/089,818B
; FILING DATE: June 3, 1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LAZAR, STEVEN R
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: GI 5290A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8260
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-089-818B-8

Query Match          52.9%; Score 37; DB 10; Length 267;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      2 THRIHWS 9
        :|||:
Db      154 SHEVHMET 161
```

```
RESULT 24
US-10-139-814-12
; Sequence 12, Application US/10139814
; Publication No. US20030134790A1
; GENERAL INFORMATION:
; APPLICANT: Langenfeld, John
; TITLE OF INVENTION: BONE MORPHOGENETIC PROTEIN-2 IN THE TREATMENT AND DIAGNOSIS OF
; FILE REFERENCE: 273/136 Michael J. Wise
; CURRENT APPLICATION NUMBER: US/10/139,814
; CURRENT FILING DATE: 2002-05-02
```



PRIOR APPLICATION NUMBER: US60/261,252  
PRIOR FILING DATE: 2001-01-12  
PRIOR APPLICATION NUMBER: US10/044,716  
PRIOR FILING DATE: 2002-01-11  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 12  
LENGTH: 267  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (361)..(741)  
OTHER INFORMATION: DAN domain  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (484)..(723)  
OTHER INFORMATION: Cysteine knot region  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (490)..(723)  
OTHER INFORMATION: C-terminal cysteine knot-like domain  
US-10-139-814-12

Query Match 52.9%; Score 37; DB 12; Length 267;  
Best Local Similarity 50.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 THRIWES 9  
DB 154 SHEVHWET 161

RESULT 25  
US-10-286-152A-52  
Sequence 52, Application US/10286152A  
Publication No. US20030134308A1  
GENERAL INFORMATION:  
APPLICANT: Alcon Research, Ltd.  
TITLE OF INVENTION: Bone Morphogenic Proteins (BMP), BMP Receptors and BMP Binding P  
TITLE OF INVENTION: and Their Use in the Diagnosis and Treatment of Glaucoma  
FILE REFERENCE: 2312 US  
CURRENT APPLICATION NUMBER: US/10/286,152A  
CURRENT FILING DATE: 2002-02-28  
NUMBER OF SEQ ID NOS: 54  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 52  
LENGTH: 267  
TYPE: PRT  
ORGANISM: homo sapiens  
US-10-286-152A-52

Query Match 52.9%; Score 37; DB 12; Length 267;  
Best Local Similarity 50.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 THRIWES 9  
DB 154 SHEVHWET 161

RESULT 26  
US-10-044-716-12  
Sequence 12, Application US/10044716  
Publication No. US2002015986A1  
GENERAL INFORMATION:  
APPLICANT: LANGENFELD, John  
TITLE OF INVENTION: BONE MORPHOGENETIC PROTEIN-2 IN THE TREATMENT AND DIAGNOSIS OF CA  
FILE REFERENCE: 270/070US  
CURRENT APPLICATION NUMBER: US/10/044,716  
CURRENT FILING DATE: 2002-01-11  
PRIOR APPLICATION NUMBER: US60/261,252

PRIOR FILING DATE: 2001-01-12  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 12  
LENGTH: 267  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (361)..(741)  
OTHER INFORMATION: DAN domain  
NAME/KEY: misc feature  
LOCATION: (484)..(723)  
OTHER INFORMATION: Cysteine knot region  
NAME/KEY: misc feature  
LOCATION: (490)..(723)  
OTHER INFORMATION: C-terminal cysteine knot-like domain  
US-10-044-716-12

Query Match 52.9%; Score 37; DB 14; Length 267;  
Best Local Similarity 50.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 THRIWES 9  
DB 154 SHEVHWET 161

RESULT 27  
US-09-887-552A-2  
Sequence 2, Application US/09887552A  
Patent No. US20020162131A1  
GENERAL INFORMATION:  
APPLICANT: Brennan, Thomas J.  
TITLE OF INVENTION: TRANSGENIC MICE CONTAINING CERBERUS GENE  
FILE REFERENCE: R-67  
CURRENT APPLICATION NUMBER: US/09/887,552A  
CURRENT FILING DATE: 2001-06-21  
PRIOR APPLICATION NUMBER: US 60/213,670  
PRIOR FILING DATE: 2000-06-21  
PRIOR APPLICATION NUMBER: US 60/266,046  
PRIOR FILING DATE: 2001-02-01  
PRIOR APPLICATION NUMBER: US 60/282,668  
PRIOR FILING DATE: 2001-04-09  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 272  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-887-552A-2

Query Match 52.9%; Score 37; DB 10; Length 272;  
Best Local Similarity 50.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 THRIWES 9  
DB 154 SHEVHWET 161

RESULT 28  
US-09-089-818B-2  
Sequence 2, Application US/09089818B  
Patent No. US20020164682A1  
GENERAL INFORMATION:  
APPLICANT: Follettie, Maximilian  
TITLE OF INVENTION: Mammalian Cerberus-Like Protein &  
TITLE OF INVENTION: Compositions  
NUMBER OF SEQUENCES: 12

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Genetics Institute, Inc.  
;; STREET: 87 CambridgePark Drive  
;; CITY: Cambridge  
;; STATE: Massachusetts  
;; COUNTRY: US  
;; ZIP: 02140  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/089,818B  
;; FILING DATE: June 3, 1998  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: LAZAR, STEVEN R  
;; REGISTRATION NUMBER: 32,618  
;; REFERENCE/DOCKET NUMBER: GI 5290A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 498-8260  
;; TELEFAX: (617) 876-5851  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 272 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-09-089-818B-2

Query Match 52.9%; Score 37; DB 10; Length 272;  
Best Local Similarity 50.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 THRIHWES 9  
:|:|:|:  
Db 154 SHEVHWET 161

RESULT 29  
US-10-106-698-4945  
; Sequence 4945, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
; FILE REFERENCE: PA005P1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 4945  
; LENGTH: 292  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (25)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: MISC FEATURE  
; LOCATION: (201)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: MISC FEATURE  
; LOCATION: (242)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-106-698-4945

Query Match 52.9%; Score 37; DB 15; Length 292;  
Best Local Similarity 50.0%; Pred. No. 2.3e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 ITHRIHWESA 10  
:|:|:|:  
Db 34 LTRVHWMAEA 43

RESULT 30  
US-10-097-340-93  
; Sequence 93, Application US/10097340  
; Publication No. US20030087250A1  
; GENERAL INFORMATION:  
; APPLICANT: John MONAHAN  
; APPLICANT: Manjula GANNAVARAPU  
; APPLICANT: Sebastian HOERSCH  
; APPLICANT: Shubhangi KAMATRAR  
; APPLICANT: Steve G. KOVATS  
; APPLICANT: Rachel E. MEYERS  
; APPLICANT: Michael MORRISSEY  
; APPLICANT: Peter OLANDT  
; APPLICANT: Ami SEN  
; APPLICANT: Peter VEIBY  
; APPLICANT: Gordon B. MILLS  
; APPLICANT: Robert C. BAST, Jr.  
; APPLICANT: Karen LU  
; APPLICANT: Rosemarie SCHMANDT  
; APPLICANT: Xumei ZHAO  
; APPLICANT: Karen GLATT  
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
; FILE REFERENCE: MRI-030  
; CURRENT APPLICATION NUMBER: US/10/097,340  
; CURRENT FILING DATE: 2002-03-14  
; PRIOR APPLICATION NUMBER: 60/276,025  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/325,149  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/276,026  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/324,967  
; PRIOR FILING DATE: 2001/09/26  
; PRIOR APPLICATION NUMBER: 60/311,732  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/325,102  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/323,580  
; PRIOR FILING DATE: 2001-09-19  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 93  
; LENGTH: 301  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-097-340-93

Query Match 52.9%; Score 37; DB 15; Length 301;  
Best Local Similarity 50.0%; Pred. No. 2.4e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 ITHRIHWESA 10  
:|:|:|:  
Db 43 LTRVHWMAEA 52

RESULT 31  
US-10-139-262-2  
; Sequence 2, Application US/10139262  
; Publication No. US20020128459A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAMURA, YUSUKE

APPLICANT: TANAKA, TOSHIHIRO  
APPLICANT: TSUKADA, SHUICHI  
TITLE OF INVENTION: PHYSIOLOGICALLY ACTIVE PROTEIN ORIGINATING IN MAMMALS  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/10/139,262  
CURRENT FILING DATE: 2002-05-07  
PRIOR APPLICATION NUMBER: US/09/380,287A  
PRIOR FILING DATE: 1999-11-16  
PRIOR APPLICATION NUMBER: JP 1997-62259  
PRIOR FILING DATE: 1997-02-28  
PRIOR APPLICATION NUMBER: JP 1998-62263  
PRIOR FILING DATE: 1998-02-25  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 385  
TYPE: PRT  
ORGANISM: Oryctolagus cuniculus  
US-10-139-262-2

Query Match 52.9%; Score 37; DB 14; Length 385;  
Best Local Similarity 41.7%; Pred. No. 3e+02; 2; Indels 0; Gaps 0;  
Matches 5; Conservative 5; Mismatches 2;

QY 1 ITRIHVESASL 12  
Db 355 LSHSLYWEVSL 366

RESULT 32  
US-10-255-969-2  
Sequence 2, Application US/10255969  
Publication No. US20030083486A1  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: TANAKA, TOSHIHIRO  
APPLICANT: TSUKADA, SHUICHI  
TITLE OF INVENTION: PHYSIOLOGICALLY ACTIVE PROTEIN ORIGINATING IN MAMMALS  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/10/255,969  
CURRENT FILING DATE: 2002-09-27  
PRIOR APPLICATION NUMBER: US/09/380,287  
PRIOR FILING DATE: 1999-11-16  
PRIOR APPLICATION NUMBER: JP 1997-62259  
PRIOR FILING DATE: 1997-02-28  
PRIOR APPLICATION NUMBER: JP 1998-62263  
PRIOR FILING DATE: 1998-02-25  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 385  
TYPE: PRT  
ORGANISM: Oryctolagus cuniculus  
US-10-255-969-2

Query Match 52.9%; Score 37; DB 15; Length 385;  
Best Local Similarity 41.7%; Pred. No. 3e+02; 2; Indels 0; Gaps 0;  
Matches 5; Conservative 5; Mismatches 2;

QY 1 ITRIHVESASL 12  
Db 355 LSHSLYWEVSL 366

RESULT 33  
US-09-764-891-4103  
Sequence 4103, Application US/09764891  
Publication No. US20030077808A1  
GENERAL INFORMATION:  
APPLICANT: ROSEN ET AL.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC006  
CURRENT APPLICATION NUMBER: US/09/764,891

CURRENT FILING DATE: 2001-01-17  
Prior Application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 10231  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4103  
LENGTH: 390  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-764-891-4103

Query Match 52.9%; Score 37; DB 11; Length 390;  
Best Local Similarity 66.7%; Pred. No. 3.1e+02; 1; Indels 0; Gaps 0;  
Matches 6; Conservative 2; Mismatches 1;

QY 1 ITRIHVES 9  
Db 333 ITRNIHYEN 341

RESULT 34  
US-10-103-313-418  
Sequence 418, Application US/10103313  
Publication No. US20030082758A1  
GENERAL INFORMATION:  
APPLICANT: ROSEN ET AL.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PZ07C1  
CURRENT APPLICATION NUMBER: US/10/103,313  
CURRENT FILING DATE: 2002-03-12  
NUMBER OF SEQ ID NOS: 653  
Prior Application removed - See File Wrapper or Palm  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 418  
LENGTH: 390  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-103-313-418

Query Match 52.9%; Score 37; DB 15; Length 390;  
Best Local Similarity 66.7%; Pred. No. 3.1e+02; 1; Indels 0; Gaps 0;  
Matches 6; Conservative 2; Mismatches 1;

QY 1 ITRIHVES 9  
Db 333 ITRNIHYEN 341

RESULT 35  
US-10-103-313-541  
Sequence 541, Application US/10103313  
Publication No. US20030082758A1  
GENERAL INFORMATION:  
APPLICANT: ROSEN ET AL.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PZ07C1  
CURRENT APPLICATION NUMBER: US/10/103,313  
CURRENT FILING DATE: 2002-03-12  
NUMBER OF SEQ ID NOS: 653  
Prior Application removed - See File Wrapper or Palm  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 541  
LENGTH: 390  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-103-313-541

Query Match 52.9%; Score 37; DB 15; Length 390;  
Best Local Similarity 66.7%; Pred. No. 3.1e+02; 1; Indels 0; Gaps 0;  
Matches 6; Conservative 2; Mismatches 1;

QY 1 ITRIHVES 9  
Db 333 ITRNIHYEN 341

```
RESULT 36
US-10-139-262-6
; Sequence 6, Application US/10139262
; Publication No. US20020128459A1
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: TANAKA, TOSHIHIRO
; APPLICANT: TSUKADA, SHUICHI
; TITLE OF INVENTION: PHYSIOLOGICALLY ACTIVE PROTEIN ORIGINATING IN MAMMALS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/139,262
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US/09/380,287A
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: JP 1997-62259
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: JP 1998-62263
; PRIOR FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 130
; OTHER INFORMATION: Xaa is Tyr or His or Gln or Asn or Lys or Asp or Glu
US-10-139-262-6

Query Match      52.9%; Score 37; DB 14; Length 433;
Best Local Similarity 41.7%; Pred. No. 3.4e+02;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Oy      1 ITHRHWESASL 12
Db      403 LSHSLYWEVSSL 414

RESULT 37
US-10-255-969-6
; Sequence 6, Application US/10255969
; Publication No. US20030083486A1
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: TANAKA, TOSHIHIRO
; APPLICANT: TSUKADA, SHUICHI
; TITLE OF INVENTION: PHYSIOLOGICALLY ACTIVE PROTEIN ORIGINATING IN MAMMALS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/255,969
; CURRENT FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: US/09/380,287
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: JP 1997-62259
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: JP 1998-62263
; PRIOR FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 130
; OTHER INFORMATION: Xaa is Tyr or His or Gln or Asn or Lys or Asp or Glu
US-10-255-969-6

Query Match      52.9%; Score 37; DB 15; Length 433;
Best Local Similarity 41.7%; Pred. No. 3.4e+02;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Oy      1 ITHRHWESASL 12
Db      403 LSHSLYWEVSSL 414
```

```
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Oy      1 ITHRHWESASL 12
Db      403 LSHSLYWEVSSL 414

RESULT 38
US-10-139-262-4
; Sequence 4, Application US/10139262
; Publication No. US20020128459A1
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: TANAKA, TOSHIHIRO
; APPLICANT: TSUKADA, SHUICHI
; TITLE OF INVENTION: PHYSIOLOGICALLY ACTIVE PROTEIN ORIGINATING IN MAMMALS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/139,262
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US/09/380,287A
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: JP 1997-62259
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: JP 1998-62263
; PRIOR FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-10-139-262-4

Query Match      52.9%; Score 37; DB 14; Length 440;
Best Local Similarity 41.7%; Pred. No. 3.4e+02;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Oy      1 ITHRHWESASL 12
Db      410 LSHSLYWEVSSL 421

RESULT 39
US-10-255-969-4
; Sequence 4, Application US/10255969
; Publication No. US20030083486A1
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: TANAKA, TOSHIHIRO
; APPLICANT: TSUKADA, SHUICHI
; TITLE OF INVENTION: PHYSIOLOGICALLY ACTIVE PROTEIN ORIGINATING IN MAMMALS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/255,969
; CURRENT FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: US/09/380,287
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: JP 1997-62259
; PRIOR FILING DATE: 1997-02-28
; PRIOR APPLICATION NUMBER: JP 1998-62263
; PRIOR FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-10-255-969-4

Query Match      52.9%; Score 37; DB 15; Length 440;
Best Local Similarity 41.7%; Pred. No. 3.4e+02;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Oy      1 ITHRHWESASL 12
```

Db 410 LSHSLYWEVSSL 421

```

RESULT 40
US-09-938-330-2
; Sequence 2, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OR INVENTION: No. US20020115838A1el Human Proteases and Polynucleotides Encodin
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 451
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-2

```

```

Query Match 52.9%; Score 37; DB 10; Length 451;
Best Local Similarity 54.5%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 HRIHESASLL 13
Db 434 HRFHMSRCSKL 444

```

Search completed: August 28, 2003, 14:22:54  
Job time : 58 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 28, 2003, 14:10:13 ; Search time 16 Seconds  
(without alignments)  
78.137 Million cell updates/sec

Title: US-09-845-738C-1  
Perfect score: 70  
Sequence: 1 ITHRHWSASL1 13

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR.76:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	100.0	1663	1 C3HU	complement C3 prec
2	49	70.0	726	2 A27602	complement C3 - ra
3	46	65.7	267	2 A82997	hypothetical prote
4	44	62.9	516	2 S67037	SMF3 protein - yea
5	42	60.0	211	2 H83239	pseudouridine synt
6	42	60.0	401	2 B82521	hypothetical prote
7	41	58.6	615	2 B86713	hypothetical prote
8	40	57.1	229	2 A13289	hypothetical cytos
9	40	57.1	259	2 T29569	hypothetical prote
10	39	55.7	406	2 T50894	hydroxyneurosporen
11	39	55.7	567	2 C69611	ABC transporter re
12	39	55.7	574	2 A81790	ABC transporter re
13	39	55.7	574	2 A81414	ABC transporter re
14	38	54.3	228	2 A12913	conserved hypotet
15	38	54.3	242	2 C70895	hypothetical prote
16	38	54.3	266	2 D97688	hypothetical prote
17	38	54.3	280	2 S46699	hypothetical prote
18	38	54.3	280	2 T21876	hypothetical prote
19	38	54.3	1896	2 T08851	Down syndrome cell
20	37.5	53.6	609	2 A80500	glutamine-fructose
21	37	52.9	220	2 S16511	probable PR1 prot
22	37	52.9	227	2 C82625	conserved hypotet
23	37	52.9	249	2 T17134	hypothetical prote
24	37	52.9	280	2 C86317	protein T1002.23
25	37	52.9	314	2 P70505	probable trna delt
26	37	52.9	341	2 T25052	hypothetical prote
27	37	52.9	421	2 F87040	conserved hypotet
28	37	52.9	698	2 I60218	hypothetical prote
29	37	52.9	851	1 WMBE09	gene UL9 protein -

30	37	52.9	858	2 T18946	probable phosphol
31	37	52.9	1455	2 G86466	hypothetical prote
32	36	51.4	147	2 G86749	conserved hypotet
33	36	51.4	164	2 S29392	estradiol-stimulat
34	36	51.4	171	2 C70542	hypothetical prote
35	36	51.4	261	2 A26344	carbonate dehydrat
36	36	51.4	291	2 A12357	hypothetical prote
37	36	51.4	300	2 C70943	hypothetical prote
38	36	51.4	320	2 S35007	modulation protein
39	36	51.4	326	2 T49666	myb-related protei
40	36	51.4	345	2 JC7681	sepin 3B - human
41	36	51.4	348	2 T21627	hypothetical prote
42	36	51.4	354	2 D41080	probable aldolase
43	36	51.4	421	2 S26605	myb-related protei
44	36	51.4	426	1 B71249	hypothetical prote
45	36	51.4	448	2 A83775	hypothetical prote

## ALIGNMENTS

## RESULT 1

C3HU  
complement C3 precursor [validated] - human  
N:Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit;  
C:Species: Homo sapiens (man)  
C:Date: 28-Aug-1985 #sequence-revision 28-Aug-1985 #text-change 08-Dec-2000  
C:Accession: A94065; A37999; A92187; A27603; A23435; A45830; B45830; A01257; A01258  
R:de Bruijn, M.H.L.; Fey, G.H.  
Proc. Natl. Acad. Sci. U.S.A. 82, 708-712, 1985  
A:Title: Human complement component C3: cDNA coding sequence and derived primary structu  
A:Reference number: A94065; MUID:85140166; PMID:2579375  
A:Accession: A94065  
A:Molecule type: mRNA  
A:Residues: 1-1663 <DEB>  
A:Cross-references: GB:X02765; NID:G179664; PID:AAA8532.1; PID:G179665  
R:Viik, D.P.; Amiguet, P.; Moffat, G.J.; Fey, M.; Amiguet-Barrias, F.; Wetzel, R.A.; Tack,  
Biochemistry 30, 1080-1085, 1991  
A:Title: Structural features of the human C3 gene: intron/exon organization, transcripti  
A:Reference number: A37999; MUID:9113687; PMID:1703437  
A:Contents: intron/exon structure of gene  
A:Accession: A37999  
A:Molecule type: DNA  
A:Residues: 1-25 <VTK>  
A:Cross-references: GB:M63423  
R:Hugli, T.E.  
J. Biol. Chem. 250, 8293-8301, 1975  
A:Title: Human anaphylatoxin (C3a) from the third component of complement.  
A:Reference number: A92187; MUID:76069169; PMID:1238393  
A:Accession: A92187  
A:Molecule type: protein  
A:Residues: 672-680, 'N', 682-699, 'Q', 701-748 <HUG>  
R:Daouadaki, M.E.; Becherer, J.D.; Lambrie, J.D.  
J. Immunol. 140, 1577-1580, 1988  
A:Title: A 34-amino acid peptide of the third component of complement mediates properdin  
A:Reference number: A27603; MUID:86154452; PMID:3279119  
A:Accession: A27603  
A:Molecule type: protein  
A:Residues: 1409-1563 <DNO>  
R:Hellman, U.; Eggertsen, G.; Engstrom, A.; Sjogvist, J.  
Biochem. J. 230, 353-361, 1985  
A:Title: Amino acid sequence of the trypsin-generated C3d fragment from human complement  
A:Reference number: A23435; MUID:86025442; PMID:3876831  
A:Accession: A23435  
A:Molecule type: protein  
A:Residues: 1002-1012, 'E', 1014-1303 <HEU>  
A:Note: sequence corresponding to residues 1072-1100 was not determined but was taken fr  
R:Poznansky, M.C.; Clossold, P.W.; Lachmann, P.J.  
J. Immunol. 143, 1254-1258, 1989  
A:Title: The difference between human C3f and C3s results from a single amino acid chang  
A:Reference number: A45830; MUID:89309808; PMID:2473125

A:Accession: A45830  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1212-1215, 'N', 1217-1223 <PO2>  
A:Note: this is the C3S allele  
A:Accession: B45830  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1212-1223 <PO2>  
A:Polymer: K.; Sottrup-Jensen, L.  
FEBS Lett. 315, 85-90, 1993  
A:Title: Disulfide bridges in human complement component C3b.  
A:Reference number: S27041, MUID:93106233; PMID:8416818  
A:Contents: annotation: disulfide bonds  
A:Comment: The sequence shown is the C3 fast (C3F) allele, which is found mainly in Caucasian alternative complement pathways, releases the C3a anaphylatoxin from the amino end of the native-complement-pathway C3/C5 convertase.  
A:Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.  
A:Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign particles, classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by products of the alternative pathway.  
A:Comment: The major site of synthesis of this plasma protein is the liver.  
A:GeneticB:  
A:Gene: GDB:C3  
A:Cross-references: GDB:119044; OMIM:120700  
A:Map position: 19p13.3-19p13.3  
A:Note: contains 41 exons  
C:Superfamily: alpha-2-macroglobulin  
C:Keywords: acute phase; complement alternate pathway; complement pathway; glycoprotein, F.1-32/Domian: signal sequence #status predicted <SIG>  
F.123-667/Product: complement C3 and C3b beta chain #status predicted <C3BB>  
F.123-667, 672-1663/Product: complement C3 #status predicted <CC3>  
F.123-667, 749-1663/Product: C3b #status predicted <C3B>  
F.1672-1663/Product: complement C3 alpha chain #status predicted <CC3A>  
F.1672-748/Product: C3a anaphylatoxin #status predicted <C3T>  
F.748-1663/Product: C3b alpha' chain #status predicted <CC3A>  
F.1948-1303/Product: C3d fragment #status predicted <CDK>  
F.1955-1303/Product: C3dg fragment #status predicted <CDG>  
F.1955-1001/Product: C3g fragment #status predicted <C3G>  
F.1002-1303/Product: C3g fragment #status experimental <CCD>  
F.1424-1457/Region: properdin binding  
F.85, 939/Binding site: carbohydrate (asn) (covalent) #status experimental  
F.555-816, 627-662, 693-720, 694-727, 707-728, 873-1513, 1101-1158, 1358-1489, 1389-1458, 1506-1515  
F.748-749/Cleavage site: Arg-Ser (C3 convertase) #status predicted  
F.1954-955/Cleavage site: Arg-Gln (complement factor I) #status predicted  
F.11010-1013/Cross-link: cholesteryl (Cys-Gln) #status experimental  
F.1303-1304/Cleavage site: Arg-Ser (complement factor I) #status predicted  
F.1350-1321/Cleavage site: Arg-Ser (complement factor I) #status predicted  
F.11617/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 100.0%; Score 70; DB 1; Length 1663;  
Best Local Similarity 100.0%; Pred. No. 0.00065;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ITHRIHWSASLL 13  
|||||||  
Db 1307 ITHRIHWSASLL 1319

RESULT 2  
A27602  
N:Complement C3 - rabbit (fragment)  
C:Species: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit;  
C:Date: 15-Dec-1988 #sequence\_revision 07-Oct-1994 #text\_change 16-Jul-1999  
A:Accession: A27602  
R:Kusano, M.; Choi, N.H.; Tomita, M.; Yamamoto, K.; Migita, S.; Sekiya, T.; Nishimura, S.; Imumoto, I. Invest. 15, 365-378, 1986  
A:Title: Nucleotide sequence of cDNA and derived amino acid sequence of rabbit complement C3.  
A:Reference number: A27602; MUID:87006507; PMID:3019881  
A:Accession: A27602  
A:Molecule type: mRNA  
A:Residues: 1-726 <KUS>

A:Cross-references: GB:3932334; NID:g164862; PIDN:AA31190.1;PID:g164863  
 C:Comment: Complement C3 contains two chains, formed by removal of four residues and 1in  
 alternative complement pathways, releases the C3a anaphylatoxin from the amino end of t  
 native-complement-pathway C3/C5 convertase.  
 C:Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.  
 C:Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign pa  
 e classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by pro  
 C:Comment: The major site of synthesis of this plasma protein is the liver.  
 C:Superfamily: alpha2-macroglobulin  
 C:Keywords: acute phase, complement alternate pathway, complement pathway, glycoprotein;

Query Match 70.0%; Score 49; DB 2; Length 726;  
 Best Local Similarity 69.2%; Pred. No. 1.3;  
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ITRHWESASL 13  
 : ||| : |||||  
 Db 370 VKRHWDSASL 382

RESULT 3  
 A82997  
 hypotheical protein PA5194 [imported] - Pseudomonas aeruginosa (strain PAO1)  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #ext\_change 31-Dec-2000  
 C:Accession: A82997  
 R:Stover, C.K.; Plam, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lartig, K.; Llm,  
 y; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A>Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pat  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: A82997  
 A>Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-267 <SNO>  
 A:Cross-references: GB:AE004932; GB:AE004091; NID:g9951493; PIDN:AA608579.1; GSPDB:GN001  
 A:Experimental source: Strain PAO1  
 C:Genetics:

A:Gene: PA5194

Query Match 65.7%; Score 46; DB 2; Length 267;  
 Best Local Similarity 58.3%; Pred. No. 1.4;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ITRHWESASL 12  
 : ||| : |||  
 Db 120 IAHHLWQHASL 131

RESULT 4  
 S67037  
 SMP3 protein - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein O3527; protein YOR149C  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #ext\_change 21-Jul-2000  
 C:Accession: S67037; S13750  
 R:Botdome, R.; Camasse, A.; Madania, A.; Martin, R.P.; Poch, O.; Tarasov, I.A.; Wince  
 submitted to the Protein Sequence Database, July 1996  
 A:Reference number: S67032  
 A:Accession: S67037  
 A:Molecule type: DNA  
 A:Residues: 1-516 <BOR>  
 A:Cross-references: EMBL:Z75057; NID:g1420374; PID:e252038; PID:g1420375; MIPS:YOR149C  
 A:Experimental source: Strain 5288C  
 R:lie, K.; Araki, H.; Oshima, Y.  
 M.Ol. Gen. Genet. 225, 257-265, 1991  
 A>Title: Mutations in a Saccharomyces cerevisiae host showing increased holding stabilit  
 A:Reference number: S13750; MUID:01172125; PMID:2005867  
 A:Accession: S13750  
 A:Molecule type: DNA  
 A:Residues: 1-121; 'IK', 124-162, 'G', 164-168, 'R', 170-278, 'L', 280-516 <IRI>  
 A:Cross-references: EMBL:X58121; NID:g4497; PIDN:CAA41123.1; PID:g4498



C:Genetics:  
A:Gene: SGD:SMP3  
A:Cross-references: SGD:S0005675; MIPS:YOR149c  
A:Map position: 15R  
C:Key words: transmembrane protein  
F:9-25/Domain: transmembrane #status predicted <TM1>  
F:189-205/Domain: transmembrane #status predicted <TM2>  
F:215-231/Domain: transmembrane #status predicted <TM3>  
F:271-287/Domain: transmembrane #status predicted <TM4>  
F:344-360/Domain: transmembrane #status predicted <TM5>

Query Match 62.9%; Score 44; DB 2; Length 516;  
Best Local Similarity 63.6%; Pred. No. 6.6;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 HRIHWSASL 13  
DB 207 YRHWKSPSL 217

## RESULT 5

pseudouridine synthase RluA PA3246 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: H83239  
R:Stover, C.K.; Pham, X.O.; Brwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: H83239  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-211 <STO>  
A:Cross-references: GB:AE004747; GB:AE004091; NID:9949362; PIDN:AAG0634.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: rluA; PA3246

Query Match 60.0%; Score 42; DB 2; Length 211;  
Best Local Similarity 46.2%; Pred. No. 5.6;  
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 ITHRHWSASL 13  
DB 52 IVHRLDWETSGIM 64

## RESULT 6

hypothetical protein XF2735 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: B82521  
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: B82521  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-401 <SIM>  
A:Cross-references: GB:AE004080; GB:AE003849; NID:99107971; PIDN:AAP8520.1; GSPDB:GN001  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, R  
as-Neto, E.; Docena, C.; El-dorri, H.; Facinani, A.F.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
J:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.B.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Mirocica, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF2735

Query Match 60.0%; Score 42; DB 2; Length 401;  
Best Local Similarity 45.5%; Pred. No. 11;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ITHRHWSAS 11  
DB 334 LAHRVHWDERS 344

## RESULT 7

hypothetical protein cydC [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 24-Aug-2001  
C:Accession: B86713  
R:Bojorin, A.; Winkler, P.; Mager, S.; Jallion, O.; Malame, K.; Weissenbach, J.; Ehrlic  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp  
A:Reference number: A86625; MUID:21235186; PMID:111337471  
A:Accession: B86713  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-615 <STO>  
A:Cross-references: GB:AE005176; PID:912723617; PIDN:AAK04804.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: cydC  
C:Superfamily: Mycobacterium tuberculosis probable ABC transporter cyd; ATP-binding cas

Query Match 58.6%; Score 41; DB 2; Length 615;  
Best Local Similarity 66.7%; Pred. No. 27;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 THRHWSA 10  
DB 509 THRLWLSS 517

## RESULT 8

hypothetical cytosolic protein BMEI0303 [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 03-Jun-2002  
C:Accession: A13289  
R:DelVecchio, V.G.; Kapral, V.; Redhar, R.J.; Parra, G.; Mijer, C.; Los, T.; Ivanova, N  
; Mazur, M.; Golsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Levesc  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A:Reference number: A13252; PMID:11756688  
A:Accession: A13289  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-229 <KUR>  
A:Cross-references: GB:AE008917; PIDN:ALU51484.1; PID:917982196; GSPDB:GN00190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI0303  
A:Map position: 1  
C:Superfamily: Rickettsia prowazekii hypothetical protein RP073  
Query Match 57.1%; Score 40; DB 2; Length 229;  
Best Local Similarity 58.3%; Pred. No. 14;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ITHRHESASL 12  
|:|:|:|:|:  
Db 137 ITRRHMSANL 148

RESULT 9  
T93569  
hypothetical protein C44C1.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
C:Accession: T93569  
R:Bradshaw, H.; Stelives, L.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of C. elegans cosmid C44C1.  
A:Reference number: Z20642  
A:Accession: T93569  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-259 <BRA>  
A:Cross-references: EMBL:U41030; PIDN:AAA82366.1; CESP:C44C1.1  
C:Genetics:  
A:Gene: CESP:C44C1.1  
A:Introns: 34/3; 82/1; 105/1; 146/2  
C:Superfamily: Caenorhabditis elegans hypothetical protein C44C1.1

Query Match 57.1%; Score 39; DB 2; Length 259;  
Best Local Similarity 75.0%; Pred. No. 16;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 THRHMS 9  
|:|:|:|:  
Db 185 THRLHMS 192

RESULT 10  
T50894  
hydroxymyosin dehydrogenase [imported] - Rubrivivax gelatinosus  
C:Species: Rubrivivax gelatinosus  
C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 21-Jul-2000  
C:Accession: T50894  
R:Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsura, K.; Shimada, K.  
submitted to the EMBL Data Library, November 1999  
A:Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photosynth  
A:Reference number: Z25270  
A:Accession: T50894  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-406 <NAG>  
A:Cross-references: EMBL:AB034704; PIDN:BAA94047.1  
A:Experimental source: strain IL144  
C:Genetics:  
A:Gene: crtc

Query Match 55.7%; Score 39; DB 2; Length 406;  
Best Local Similarity 60.0%; Pred. No. 39;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 RIHWESASL 13  
|:|:|:|:  
Db 192 RIHWDESL 201

RESULT 11  
C69611  
ABC transporter required for expression of cytochrome bd (ATP-) cydC - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 02-Feb-2001  
C:Accession: C69611  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
C.; Bron, S.; Brulle, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Gallier  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
A:Authors: Lauber, U.J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portecelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, B.; Sadale, Y.; Sato, T.; Scanlon  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosa, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zundstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: C69611  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-567 <KUN>  
A:Cross-references: GB:Z99123; GB:AL009126; NID:G2636240; PIDN:CAB15900.1; PID:ell186373;  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: cydC  
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology  
C:Keywords: ATP; nucleotide binding; P-loop  
F:343-537/Domain: ATP-binding cassette homology <ABC>  
F:360-367/Region: nucleotide-binding motif A (P-loop)

Query Match 55.7%; Score 39; DB 2; Length 567;  
Best Local Similarity 83.3%; Pred. No. 56;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 THRHMS 7  
|:|:|:|:  
Db 519 THRLHM 524

RESULT 12  
AB1790  
ABC transporter required for expression of cytochrome BD homolog cydC [imported] - Listeria  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
C:Accession: AB1790  
R:Glaser, P.; Franke, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker  
J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fsthi, H.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurajkar, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schueter, T.; Simoes, N.; Tisser, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AB1790  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-574 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CAC98090.1; PID:G16415399; GSPDB:GN00178  
A:Experimental source: strain C1p11262  
C:Genetics:  
A:Gene: cydC  
C:Superfamily: Mycobacterium tuberculosis probable ABC transporter cydC; ATP-binding cas

Query Match 55.7%; Score 39; DB 2; Length 574;  
Best Local Similarity 83.3%; Pred. No. 56;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 THRHMS 7  
|:|:|:|:  
Db 527 THRLHM 532

RESULT 13  
AC1414  
ABC transporter required for expression of cytochrome BD homolog cydC [imported] - Listeria  
C:Species: Listeria monocytogenes  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001

C:Accession: AC1414  
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fehli, H.; Jones, L.M.; Kars, U.  
 Science 294, 849-852, 2001  
 A:Authors: Krell, U.; Kunz, M.; Kuntz, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tietter, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of *Listeria species*.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AC1414  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-574 <GLA>  
 A:Cross-references: GB:NC 003210; PIDN:CAD00929.1; PID:g16412216; GSPDB:GN00177  
 A:Experimental source: strain EGD-e  
 C:Genetics:  
 A:Gene: cydC  
 C:Superfamily: Mycobacterium tuberculosis probable ABC transporter cydC; ATP-binding cas

Query Match 55.7%; Score 39; DB 2; Length 574;  
 Best Local Similarity 83.3%; Pred. No. 56;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 THRIHW 7  
 |||:  
 |||:  
 Db 527 THRIHW 532

RESULT 14  
 A12913  
 conserved hypothetical protein Atut2746 [imported] - Agrobacterium tumefaciens (strain C5  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
 C:Accession: A12913  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Kreppan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; MUID:21608550; PMID:11743193  
 A:Accession: A12913  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-228 <KUR>  
 A:Cross-references: GB:AE006688; PIDN:AAL43727.1; PID:g17741259; GSPDB:GN00186  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atut2746  
 A:Map position: circular chromosome  
 C:Superfamily: Rickettsia prowazekii hypothetical protein RP073

Query Match 54.3%; Score 38; DB 2; Length 228;  
 Best Local Similarity 58.3%; Pred. No. 31;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 ITHRIHWESASL 12  
 |||:  
 |||:  
 Db 131 ITRTHWNSANL 142

RESULT 15  
 C70885  
 hypothetical protein RV1085C - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C:Accession: C70885  
 R:Cole, S.T.; Broch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S  
 ; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.  
 Rajandram, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: C70895  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-242 <COL>  
 A:Cross-references: GB:AL021897; GB:AL123456; NID:g3256022; PIDN:CAA17201.1; PID:g289672  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: RV1085C  
 C:Superfamily: hemolysin III YPIQ

Query Match 54.3%; Score 38; DB 2; Length 242;  
 Best Local Similarity 55.6%; Pred. No. 33;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 3 HRHWESAS 11  
 |||:  
 |||:  
 Db 89 HRVWKSAT 97

RESULT 16  
 D97688  
 hypothetical protein AGR C 4981 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
 C:Accession: D97688  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A:Title: Genome sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A:Reference number: A97359; MUID:21608551; PMID:11743194  
 A:Accession: D97688  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-266 <KUR>  
 A:Cross-references: GB:AE007869; PIDN:AAK8461.1; PID:g15157964; GSPDB:GN00169  
 C:Genetics:  
 A:Gene: AGR C 4981  
 A:Map position: circular chromosome  
 C:Superfamily: Rickettsia prowazekii hypothetical protein RP073

Query Match 54.3%; Score 38; DB 2; Length 266;  
 Best Local Similarity 58.3%; Pred. No. 36;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 ITHRIHWESASL 12  
 |||:  
 |||:  
 Db 169 ITRTHWNSANL 180

RESULT 17  
 S46699  
 hypothetical protein YHR067W - Yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein H8025.4  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 19-Apr-2002  
 C:Accession: S46699  
 R:Latreille, P.  
 submitted to the EMBL Data Library, May 1994  
 A:Description: The sequence of *S. cerevisiae* cosmid 8025.  
 A:Reference number: S46696  
 A:Accession: S46699  
 A:Molecule type: DNA  
 A:Residues: 1-280 <LAT>  
 A:Cross-references: EMBL:U00061; NID:g487943; PID:g487947; GSPDB:GN00008; MIPS:YHR067W  
 C:Genetics:  
 A:Gene: MIPS:YHR067W  
 A:Cross-references: SGD:S0001109  
 A:Map position: 8R  
 C:Superfamily: Saccharomyces hypothetical protein YHR067W

Query Match 54.3%; Score 38; DB 2; Length 280;

Best Local Similarity 83.3%; Pred. No. 39;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 3 HRIHME 8  
DB 188 HRIHMD 193

## RESULT 18

T21876

hypothetical protein F36G9.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T21876

R:Wall, M.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19481

A:Accession: T21876

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-280 &lt;MIL&gt;

A:Cross-references: EMBL:Z81533; PIDN:CA04338.1; GSPDB:GN00023; CESP:F36G9.7

A:Experimental source: clone F36G9

C:Genetics:

A:Gene: CESP:F36G9.7

A:Map position: 5

A:Introns: 29/3; 41/1; 154/2; 197/1; 267/1

## Query Match

Best Local Similarity 54.3%; Score 38; DB 2; Length 280;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 IHMESASLL 13  
DB 7 VHMESLTL 15

## RESULT 19

T08851

Down syndrome cell adhesion protein 1 - human (fragment)

N:Alternate names: Down syndrome cell adhesion molecule

C:Species: Homo sapiens (man)

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 11-Jun-1999

C:Accession: T08851

R:Yamakawa, K.; Huo, Y.K.; Haendel, M.A.; Hubert, R.; Chen, X.N.; Lyons, G.E.; Korenberg

submitted to the EMBL Data Library, September 1997

A:Description: DSCAM: A novel member of the immunoglobulin superfamily maps in a down sy

A:Reference number: Z16495

A:Accession: T08851

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1896 &lt;YAM&gt;

A:Cross-references: EMBL:AF023449; NID:g3169765; PID:g3169766

A:Experimental source: brain; developmental stage: 14 weeks; fetal

C:Genetics:

A:Gene: DSCAM

A:Map position: 21622

A&gt;Note: derived from alternately-spliced mRNA

C:Function:

A:Description: involved in nervous system development

C:Keywords: alternative splicing

Query Match 54.3%; Score 38; DB 2; Length 1896;  
Best Local Similarity 45.5%; Pred. No. 3.1e+02;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ITHRIHMEAS 11  
DB 1688 VHTVHYGQSV 1698

## RESULT 20

AB0500

glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [imported] - Yer  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 03-Jun-2002  
C:Accession: AB0500  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titchell, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Rarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AB0500  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-609 <KUP>  
A:Cross-references: GB:AL590842; PIDN:CA093567.1; PID:g15982007; GSPDB:GN00175  
C:Genetics:

A:Gene: glms

C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)

C:Keywords: aminotransferase; intramolecular oxidoreductase; isomerase

## Query Match

Best Local Similarity 53.6%; Score 37.5; DB 2; Length 609;

Matches 8; Conservative 1; Mismatches 4; Indels 3; Gaps 1;

QY 1 ITHRIHME--SASLL 13  
DB 128 IAHVHWEQQGGSL 143

## RESULT 21

S16511

probable PRY1 protein - yeast (Pichia angusta)

C:Species: Pichia angusta

C:Date: 03-May-1994 #sequence\_revision 20-Feb-1995 #text\_change 20-Apr-2000

C:Accession: S16511

R:Brinlenderg, P.G.; Evers, M.; Waterham, H.R.; Kuipers, J.; Arnerberg, A.C.; Ab, G.

Biochim. Biophys. Acta 1008, 157-167, 1989

A:Title: Cloning and sequencing of the peroxisomal amine oxidase gene from Hansenula pol-

A:Reference number: S04963; MUID:89287321; PMID:2500147

A:Accession: S16511

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-220 &lt;BRU&gt;

A:Cross-references: EMBL:X15111; NID:g2772; PIDN:CAA33208.1; PID:g2773

A:Experimental source: strain CBS 4732

C:Genetics:

A:Gene: PRT1

## Query Match

Best Local Similarity 52.9%; Score 37; DB 2; Length 220;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 ITHRIHMEASLL 12  
DB 184 IEPRAHMTSASM 195

## RESULT 22

C62625

conserved hypothetical protein XF1894 [imported] - Xylella fastidiosa (strain 9A5C)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C:Accession: C62625

R:Anonymous. The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen-

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A62515; MUID:20365717; PMID:10910347

A&gt;Note: for a complete list of authors see reference number A59328 below

A:Accession: C62625

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-227 &lt;SIM&gt;

A:Cross-references: GB:AE004009; GB:AE003849; NID:g9106980; PIDN:AAF94700.1; GSPDB:GN001

A:Experimental source: strain 9a5c  
R:Simpon, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; B  
Brienes, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, H  
as-Melo, E.; Docena, C.; El-Dorri, H.; Facinanti, A.F.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Perro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitzjima, J.P.; Klieger, J.E.; Kuramae, E.E.; Laig  
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Marukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.U.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv  
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XP1894

Query Match 52.9%; Score 37; DB 2; Length 227;  
Best Local Similarity 54.5%; Pred. No. 46;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 HRIHWSASL 13  
Db 141 HRNHPNMLAL 151

RESULT 23  
T17134  
hypothetical protein T30A10.110 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T17134  
R:Bayan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Be  
submitted to the Protein Sequence Database, September 1999  
A:Reference number: Z18708  
A:Accession: T17134  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-249 <BEV>  
A:Cross-references: EMBL:AL117386  
A:Experimental source: cultivar Columbia; BAC clone T30A10  
C:Genetics:  
A:Map position: 4  
A:Intons: 109/2  
A:Note: T30A10.110

Query Match 52.9%; Score 37; DB 2; Length 249;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RHHMES 9  
Db 62 RHMHES 67

RESULT 24  
C86317  
protein T10022.23 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: C86317  
R:Rheologis, A.; Becker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: C86317  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-280 <STO>  
A:Cross-references: GB:AE005172; NID:g8671774; PIDN:AAF78380.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: T10022.23  
A:Map position: 1

Query Match 52.9%; Score 37; DB 2; Length 280;  
Best Local Similarity 50.0%; Pred. No. 58;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 THRIHWSASL 13  
Db 126 THRIHWSASL 137

RESULT 25  
F70505  
probable trna delc - Mycobacterium tuberculosis (strain H37Rv)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: F70505  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S  
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squires, R.; Sulterson, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:9825987; PMID:9634230  
A:Accession: F70505  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-314 <COL>  
A:Cross-references: GB:Z98209; GB:AL123456; NID:g3261838; PIDN:CAB10903.1; PID:g2292961  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: miaA  
C:Superfamily: delta(2)-isopentenylpyrophosphate transferase

Query Match 52.9%; Score 37; DB 2; Length 314;  
Best Local Similarity 50.0%; Pred. No. 65;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 HRIHWSASL 12  
Db 284 HRVHMDAGV 293

RESULT 26  
T25052  
hypothetical protein T21B4.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T25052  
R:Smyle, R.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z19974  
A:Accession: T25052  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-341 <WIL>  
A:Cross-references: EMBL:Z61124; PIDN:CAB03373.1; GSPDB:GN00020; CESP:T21B4.5  
A:Experimental source: clone T21B4  
C:Genetics:  
A:Gene: CESP:T21B4.5  
A:Map position: 2  
A:Intons: 120/1; 183/3  
C:Superfamily: Caenorhabditis hypothetical protein C49G7.2

Query Match 52.9%; Score 37; DB 2; Length 341;  
Best Local Similarity 66.7%; Pred. No. 72;



Db 226 HHMDVASASITL 238

RESULT 31  
G86466  
hypothetical protein F23M19.8 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
C/Accession: G86466  
R/Theologian, A.; Ecker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.E.; Hughes, B.; Hutzar, L.  
Nature 408: 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maltli, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, C.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: G86466  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-1456 <STO>  
A/Cross-references: GB:AE05172; NID:95091620; PIDN:AMD39608.1; GSPDB:GN00141  
C/Genetics:  
A/Map position: 1

Query Match 52.9%; Score 37; DB 2; Length 1456;  
Best Local Similarity 53.8%; Pred. No. 3.5e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 1 ITHRHESASITL 13  
|:|:|:|:|:  
Db 164 TVARLCWEASML 196

RESULT 32  
G86749  
conserved hypothetical protein ykce [imported] - Lactococcus lactis subsp. lactis (strain  
C/Species: Lactococcus lactis subsp. lactis  
C/Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C/Accession: G86749  
R/Bolotin, A.; Wincker, P.; Manger, S.; Jallion, O.; Malarne, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
A/Reference number: A86625; MUID:21235186; PMID:11337471  
A/Accession: G86749  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-147 <STO>  
A/Cross-references: GB:AE05176; PID:912723946; PIDN:AAK05097.1; GSPDB:GN00146  
C/Genetics:  
A/Experimental source: strain IL1403  
C/Superfamily: hypothetical protein yuid

Query Match 51.4%; Score 36; DB 2; Length 147;  
Best Local Similarity 71.4%; Pred. No. 43;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 THRIHWE 8  
|:|:|:|:  
Db 35 THRIHWE 41

RESULT 33  
S29392  
estradiol-stimulated protein ESPI - rat  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 23-Feb-1994 #sequence\_revision 01-Sep-1995 #text\_change 07-May-1999  
C/Accession: S29392  
R/Malik, P.; Panayotova-Helermann, M.; Pongse, O.

Mol. Cell. Endocrinol. 62, 235-242, 1989  
A/Title: Characterization of an estradiol-stimulated mRNA in the brain of adult male rats  
A/Reference number: S29392; MUID:89306157; PMID:2744228  
A/Accession: S29392  
A/Molecule type: mRNA  
A/Residues: 1-164 <NAL>  
C/Superfamily: LIM metal-binding repeat homology  
F/82-113/Domain: LIM metal-binding repeat homology <LIM>

Query Match 51.4%; Score 36; DB 2; Length 164;  
Best Local Similarity 50.0%; Pred. No. 48;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ITHRIHWE 8  
|:|:|:|:  
Db 71 LCHHWE 78

RESULT 34  
C70542  
hypothetical protein RV1585c - Mycobacterium tuberculosis (strain H37RV)  
C/Species: Mycobacterium tuberculosis  
C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C/Accession: C70542  
R/Conor, R.; Davies, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
A/Authors: Sgarbi, R.; Sultoni, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A/Reference number: A70500; MUID:98295987; PMID:9634230  
A/Accession: C70542  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-171 <COL>  
A/Cross-references: GB:Z95586; GB:AL123456; NID:93261785; PIDN:CAB09084.1; PID:92117252  
A/Experimental source: strain H37RV  
C/Genetics:  
A/Gene: RV1585c  
C/Superfamily: Mycobacterium tuberculosis hypothetical protein RV1585c

Query Match 51.4%; Score 36; DB 2; Length 171;  
Best Local Similarity 66.7%; Pred. No. 50;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 4 RHMHESASITL 12  
|:|:|:|:|:  
Db 82 RLHWEFACL 90

RESULT 35  
A26344  
carbonate dehydratase (EC 4.2.1.1) I - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 21-May-1988 #sequence\_revision 21-May-1988 #text\_change 22-Jun-1999  
C/Accession: A26344; I49573  
R/Frazer, P.D.; Curtis, P.J.  
J. Mol. Evol. 23, 294-299, 1986  
A/Title: Molecular evolution of the carbonic anhydrase genes: calculation of divergence  
A/Reference number: A26344; MUID:87169766; PMID:3104601  
A/Accession: A26344  
A/Molecule type: mRNA  
A/Residues: 1-261 <FRA>  
R/Frazer, P.; Cummings, P.; Curtis, P.  
Mol. Cell. Biol. 9, 3308-3313, 1989  
A/Title: The mouse carbonic anhydrase I gene contains two tissue-specific promoters.  
A/Accession: I49573  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-82, 'G', 83-100, 102-261 <RES>  
A/Cross-references: GB:I36655; NID:9556293; PIDN:AAA50291.1; PID:9556295  
C/Genetics:

A;Gene: Carl  
A;Intons: 13/1; 79/1; 118/3; 150/3; 171/3; 223/3  
C;Superfamily: carbonic dehydratase; carbonic anhydrase homology  
C;Keywords: carbon-oxygen lyase; hydro-lyase; zinc  
F;6-261/Domain: carbonic anhydrase homology <CAH>

Query Match 51.4%; Score 36; DB 2; Length 261;  
Best Local Similarity 62.5%; Pred. No. 80;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 HRIHWSA 10  
Db 120 HLHWMNSA 127

RESULT 36  
A;Intons: 13/1; 79/1; 118/3; 150/3; 171/3; 223/3  
C;Superfamily: carbonic dehydratase; carbonic anhydrase homology  
C;Keywords: carbon-oxygen lyase; hydro-lyase; zinc  
F;6-261/Domain: carbonic anhydrase homology <CAH>

Query Match 51.4%; Score 36; DB 2; Length 261;  
Best Local Similarity 62.5%; Pred. No. 80;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 HRIHWSA 10  
Db 120 HLHWMNSA 127

RESULT 36  
A;Intons: 13/1; 79/1; 118/3; 150/3; 171/3; 223/3  
C;Superfamily: carbonic dehydratase; carbonic anhydrase homology  
C;Keywords: carbon-oxygen lyase; hydro-lyase; zinc  
F;6-261/Domain: carbonic anhydrase homology <CAH>

Query Match 51.4%; Score 36; DB 2; Length 261;  
Best Local Similarity 62.5%; Pred. No. 80;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 HRIHWSA 10  
Db 120 HLHWMNSA 127

RESULT 36  
A;Intons: 13/1; 79/1; 118/3; 150/3; 171/3; 223/3  
C;Superfamily: carbonic dehydratase; carbonic anhydrase homology  
C;Keywords: carbon-oxygen lyase; hydro-lyase; zinc  
F;6-261/Domain: carbonic anhydrase homology <CAH>

Query Match 51.4%; Score 36; DB 2; Length 261;  
Best Local Similarity 62.5%; Pred. No. 80;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 HRIHWSA 10  
Db 120 HLHWMNSA 127

RESULT 36  
A;Intons: 13/1; 79/1; 118/3; 150/3; 171/3; 223/3  
C;Superfamily: carbonic dehydratase; carbonic anhydrase homology  
C;Keywords: carbon-oxygen lyase; hydro-lyase; zinc  
F;6-261/Domain: carbonic anhydrase homology <CAH>

Query Match 51.4%; Score 36; DB 2; Length 261;  
Best Local Similarity 62.5%; Pred. No. 80;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 HRIHWSA 10  
Db 120 HLHWMNSA 127

RESULT 36  
A;Intons: 13/1; 79/1; 118/3; 150/3; 171/3; 223/3  
C;Superfamily: carbonic dehydratase; carbonic anhydrase homology  
C;Keywords: carbon-oxygen lyase; hydro-lyase; zinc  
F;6-261/Domain: carbonic anhydrase homology <CAH>

Query Match 51.4%; Score 36; DB 2; Length 261;  
Best Local Similarity 62.5%; Pred. No. 80;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ITHRIHWSA 10

Db 287 ITHRRSWEAS 236

RESULT 38  
A;Intons: 13/1; 79/1; 118/3; 150/3; 171/3; 223/3  
C;Superfamily: carbonic dehydratase; carbonic anhydrase homology  
C;Keywords: carbon-oxygen lyase; hydro-lyase; zinc  
F;6-261/Domain: carbonic anhydrase homology <CAH>

Query Match 51.4%; Score 36; DB 2; Length 320;  
Best Local Similarity 58.3%; Pred. No. 1e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 THRIHWSASL 13  
Db 222 TVRAEWETALL 233

RESULT 39  
A;Intons: 13/1; 79/1; 118/3; 150/3; 171/3; 223/3  
C;Superfamily: carbonic dehydratase; carbonic anhydrase homology  
C;Keywords: carbon-oxygen lyase; hydro-lyase; zinc  
F;6-261/Domain: carbonic anhydrase homology <CAH>

Query Match 51.4%; Score 36; DB 2; Length 320;  
Best Local Similarity 58.3%; Pred. No. 1e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 THRIHWSASL 13  
Db 222 TVRAEWETALL 233

RESULT 39  
A;Intons: 13/1; 79/1; 118/3; 150/3; 171/3; 223/3  
C;Superfamily: carbonic dehydratase; carbonic anhydrase homology  
C;Keywords: carbon-oxygen lyase; hydro-lyase; zinc  
F;6-261/Domain: carbonic anhydrase homology <CAH>

Query Match 51.4%; Score 36; DB 2; Length 320;  
Best Local Similarity 58.3%; Pred. No. 1e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 THRIHWSASL 13  
Db 222 TVRAEWETALL 233

RESULT 39  
A;Intons: 13/1; 79/1; 118/3; 150/3; 171/3; 223/3  
C;Superfamily: carbonic dehydratase; carbonic anhydrase homology  
C;Keywords: carbon-oxygen lyase; hydro-lyase; zinc  
F;6-261/Domain: carbonic anhydrase homology <CAH>

Query Match 51.4%; Score 36; DB 2; Length 320;  
Best Local Similarity 58.3%; Pred. No. 1e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 THRIHWSASL 13  
Db 222 TVRAEWETALL 233

RESULT 39  
A;Intons: 13/1; 79/1; 118/3; 150/3; 171/3; 223/3  
C;Superfamily: carbonic dehydratase; carbonic anhydrase homology  
C;Keywords: carbon-oxygen lyase; hydro-lyase; zinc  
F;6-261/Domain: carbonic anhydrase homology <CAH>

Query Match 51.4%; Score 36; DB 2; Length 320;  
Best Local Similarity 58.3%; Pred. No. 1e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ITHRIHWSASL 12  
Db 146 LSHTRAQWESARL 157

RESULT 40  
A;Intons: 13/1; 79/1; 118/3; 150/3; 171/3; 223/3  
C;Superfamily: carbonic dehydratase; carbonic anhydrase homology  
C;Keywords: carbon-oxygen lyase; hydro-lyase; zinc  
F;6-261/Domain: carbonic anhydrase homology <CAH>

Query Match 51.4%; Score 36; DB 2; Length 320;  
Best Local Similarity 58.3%; Pred. No. 1e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ITHRIHWSASL 12  
Db 146 LSHTRAQWESARL 157



A/Title: Human septin 3 on chromosome 22q13.2 is upregulated by neuronal differentiation  
A/Reference number: JC7681; MUID:21222847; PMID:11322766  
A/Accession: JC7681  
A/Molecule type: mRNA  
A/Residues: 1-345 <MET>  
A/Cross-references: GB:AF285109  
C/Comment: This protein, a neuronal septin, has roles in synaptic transmission neurons.  
C/Genetics:  
A/Gene: sep3B  
A/Map position: 22q13.2

Query March 51.4%; Score 36; DB 2; Length 345;  
Best Local Similarity 55.6%; Pred. No. 1.1e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ITHRIHWES 9  
: || ||:  
Db 302 VTHNIHYET 310

Search completed: August 28, 2003, 14:11:09  
Job time : 18 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2003, 14:10:13 ; Search time 24 Seconds

(without alignments)  
25.473 Million cell updates/sec

Title: US-09-845-738C-1  
Perfect score: 70  
Sequence: 1 ITHRHVESASLL 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	70	100.0	1663	1 CO3_HUMAN	P01024 homo sapien
2	49	70.0	726	1 CO3_RABIT	P12247 cyclophilin
3	44	62.9	516	1 SMP3_YEAST	O04174 saccharomyc
4	40	57.1	396	1 RT09_HUMAN	P82933 homo sapien
5	39	55.7	336	1 PRX2_PSEST	O69054 pseudomonas
6	39	55.7	567	1 CYDC_BACCU	P94366 bacillus su
7	38	54.3	242	1 YA85_MYCTU	O53433 mycobacteri
8	38	54.3	280	1 GEM2_HUMAN	O14893 homo sapien
9	38	54.3	280	1 YHM7_YEAST	P38790 saccharomyc
10	38	54.3	611	1 GIMS_WIGBR	O6d310 w. glucosami
11	38	54.3	2012	1 DSCA_HUMAN	O60468 homo sapien
12	37.5	53.6	608	1 GIMS_YERPE	O62968 y. glucosami
13	37	52.9	220	1 PRT1_PICAN	P12806 pichia angu
14	37	52.9	269	1 GEM2_RAT	O92921 rattus norv
15	37	52.9	314	1 MIRA_MYCTU	O33233 mycobacteri
16	37	52.9	587	1 T9S3_MOUSE	O92920 mus musculu
17	37	52.9	589	1 T9S3_HUMAN	O9hd43 homo sapien
18	37	52.9	698	1 TNPX_ECOLI	O00042 escherichia
19	37	52.9	851	1 OBP_HSV11	P10193 herpes simp
20	37	52.9	1223	1 AT14_HUMAN	O6w868 homo sapien
21	36	51.4	175	1 RMP2_HUMAN	O60895 homo sapien
22	36	51.4	260	1 CAH1_MOUSE	P13614 mus musculu
23	36	51.4	260	1 CAH1_SHEEP	P48282 ovis aries
24	36	51.4	320	1 NOD1_AZOCA	O07756 azorhizobiu
25	36	51.4	345	1 SEP3_HUMAN	O9uh03 homo sapien
26	36	51.4	354	1 ALF2_RHOSH	P39271 rhodobacter
27	36	51.4	465	1 SEP3_MOUSE	O92165 mus musculu
28	36	51.4	518	1 GGT_SYNY3	P74181 synechocyst
29	36	51.4	569	1 U171_HUMAN	O12980 homo sapien
30	36	51.4	607	1 GIMS_BUCAP	O8ka75 b. glucosami
31	36	51.4	608	1 GIMS_BUCAI	P57138 b. glucosami
32	36	51.4	625	1 T9S4_HUMAN	O92544 homo sapien
33	36	51.4	843	1 VAV_RAT	P54100 rattus norv

34	36	51.4	845	1 VAV_MOUSE	P27870 mus musculu
35	36	51.4	847	1 VAV3_HUMAN	O9uk44 homo sapien
36	36	51.4	847	1 VAV3_MOUSE	O9r008 mus musculu
37	36	51.4	1663	1 CO3_RAT	P01026 rattus norv
38	36	51.4	3412	1 POLG_TBSEV	P07720 t. genome po
39	36	50.0	92	1 WIT1_HUMAN	O06250 homo sapien
40	35	50.0	99	1 YW71_MYCTU	O50692 mycobacteri
41	35	50.0	127	1 COPC_XANCU	O56797 xanthomonas
42	35	50.0	144	1 SY25_MOUSE	O35903 mus musculu
43	35	50.0	167	1 EPA3_MOUSE	O08545 mus musculu
44	35	50.0	238	1 EPA3_HUMAN	P52797 homo sapien
45	35	50.0	260	1 CAH1_MACMU	P00916 macaca mula

## ALIGNMENTS

RESULT 1  
ID CO3\_HUMAN STANDARD; PRT; 1663 AA.  
AC P01024;  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Complement C3 precursor [Contains: C3a anaphylatoxin].  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE=85140166; PubMed=2579379;  
RA "de Bruijn M.H.L., Fey G.H.;"  
RT "Human complement component C3: cDNA coding sequence and derived  
RT primary structure.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:708-712 (1985).  
RN [2]  
RP MEDLINE=672-748.  
RX MEDLINE=76069169; PubMed=1238393;  
RA Hugli T.E.;"  
RT "Human anaphylatoxin (C3a) from the third component of complement.  
RT Primary structure.";  
RL J. Biol. Chem. 250:8293-8301 (1975).  
RN [3]  
RP SEQUENCE OF 955-966, AND SUBUNITS.  
RC TISSUE=Serum;  
RX MEDLINE=95293954; PubMed=7539791;  
RA Oxvig C., Haaning J., Kristensen L., Wagner J.M., Rubin I.,  
RT "Identification of angiotensinogen and complement C3dg as novel  
RT proteins binding the proform of eosinophil major basic protein in  
RT human pregnancy serum and plasma.";  
RL J. Biol. Chem. 270:13645-13651 (1995).  
RN [4]  
RP SEQUENCE OF 988-1036.  
RX MEDLINE=82174534; PubMed=6175959;  
RA Thomas M.L., Janatova J., Gray W.R., Tack B.F.;"  
RT "Third component of human complement: localization of the internal  
RT thiolester bond.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:1054-1058 (1982).  
RN [5]  
RP SEQUENCE OF 1409-1563.  
RX MEDLINE=88154452; PubMed=3279119;  
RA Douvrouki M.E., Becherer J.D., Lambiris J.D.;"  
RT "A 34-amino acid peptide of the third component of complement  
RT mediates properdin binding.";  
RL J. Immunol. 140:1577-1580 (1988).  
RN [6]  
RP STRUCTURE BY NMR OF C3A.  
RX MEDLINE=88276894; PubMed=3260670;  
RA Nettesheim D.G., Edalji R.P., Mollison K.W., Greer J.,  
RA Zuidweg E.R.P.;"

"Secondary structure of complement component C3a anaphylatoxin in solution as determined by NMR spectroscopy: differences between RT crystal and solution conformations.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 85:5036-5040(1988).  
 RN [17]  
 RP MUTAGENESIS OF THIOESTER BOND REGION.  
 RX MEDLINE=92250565; PubMed=1577777;  
 RA Isaac L., Iseman D.E.;  
 RT "Structural requirements for thioester bond formation in human RT complement component C3. Reassessment of the role of thioester bond integrity on the conformation of C3.";  
 RT J. Biol. Chem. 267:10062-10069(1992).  
 RN [8]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=93106233; PubMed=8416818;  
 RA Dolmer K., Sottrup-Jensen L.;  
 RT "Disulfide bridges in human complement component C3b.";  
 RT FEBS Lett. 315:85-90(1993).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 996-1303.  
 RX MEDLINE=98259089; PubMed=9596584;  
 RA Nagar B., Jones R.G., Diefenbach R.J., Iseman D.E., Rini J.M.;  
 RT "X-ray crystal structure of C3d: a C3 fragment and ligand for RT complement receptor 2.";  
 RT Science 280:1277-1281(1998).  
 RN [10]  
 RP VARIANT C3F/S.  
 RX MEDLINE=89309808; PubMed=2473125;  
 RA Poznansky M.C., Clissold P.M., Lachmann P.J.;  
 RT "The difference between human C3F and C3S results from a single amino RT acid change from an asparagine to an aspartate residue at position RT 1216 on the alpha-chain of the complement component, C3.";  
 RN J. Immunol. 143:1254-1258(1989).  
 RN [11]  
 RP ERRATUM (RETRACTION).  
 RX MEDLINE=90063087; PubMed=2584723;  
 RA Poznansky M.C., Clissold P.M., Lachmann P.J.;  
 RT J. Immunol. 143:3860-3862(1989).  
 RN [12]  
 RP VARIANTS GLY-102 AND PRO-314.  
 RX MEDLINE=91011240; PubMed=1976733;  
 RA Boto M., Yong Fong K., So A.K., Koch C., Walport M.J.;  
 RT "Molecular basis of polymorphisms of human complement component C3.";  
 RN J. Exp. Med. 172:1011-1017(1990).  
 RN [13]  
 RP VARIANT ASN-549.  
 RX MEDLINE=95050640; PubMed=7961791;  
 RA Singer L., Whitehead W.T., Akama H., Katz Y., Fishelson Z., Weissel R.A.;  
 RT "Inherited human complement C3 deficiency. An amino acid substitution in the beta-chain (ASP549 to ASN) impairs C3 secretion.";  
 RN J. Biol. Chem. 269:28494-28499(1994).  
 RN [14]  
 RP VARIANT GLN-1320.  
 RA Watanabe Y., Matsumi N., Yan K., Nishimukai H., Tokunaga K., RA Juli T., Kobayashi N., Kohsaka T.;  
 RT "A novel C3 allotype C3/F02 has an amino acid substitution that may RT inhibit C3b synthesis and cause C3-hypocomplementemia.";  
 RN Mol. Immunol. 30:62-62(1993).  
 CC -1- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE CC COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL CC REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS. CC AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE CC THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES. CC -1- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C3, CC C3A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESSES. IT CC INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR CC PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND CC BASOPHILIC LEUKOCYTES.  
 CC -1- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg CC residues, forming two chains, beta and alpha, linked by a CC disulfide bond. C3 convertase activates C3 by cleaving the alpha CC chain, releasing C3a anaphylatoxin and generating C3b (beta chain

CC + alpha' chain). During pregnancy, C3dg exists as a complex CC (probably a 2:2:2 heterohexamer) with AGT and the proform of PRG2. CC POLYMORPHISMS: THERE ARE TWO ALLELES: C3S (C3 SLOW), THE MOST CC COMMON ALLELE IN ALL RACES AND C3F (C3 FAST), RELATIVELY FREQUENT CC IN CAUCASIIDS, LESS COMMON IN BLACK AMERICAN, EXTREMELY RARE IN CC ORIENTALS.  
 CC -1- DISEASE: Defects in C3 are the cause of C3 deficiency CC (MIM:120700), that can result in susceptibility to pyogenic CC infection.  
 CC -1- MISCELLANEOUS: C3B IS RAPIDLY SPLIT IN TWO POSITIONS BY FACTOR I CC AND A COFACTOR TO FORM ICB (INACTIVATED C3B) AND C3F WHICH IS CC RELEASED.  
 CC -1- MISCELLANEOUS: IC3B IS THE SLOWLY CLEAVED (POSSIBLY BY FACTOR I) CC TO FORM C3C AND C3DG. OTHER PROTEASES PRODUCE OTHER FRAGMENTS SUCH CC AS C3D OR C3G.  
 CC -1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.  
 CC -1- SIMILARITY: Contains 1 anaphylatoxin-like domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration CC between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its CC use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: K02765; AAA85332.1; -;  
 CC PIR: A94065; C3HU.  
 CC PDB: 1C3D; 18-NOV-98.  
 CC PD9; 1GHQ; 06-JUN-01.  
 CC DR SWISS-2DPAGE; P01024; HUMAN.  
 CC DR Sienaa-2DPAGE; P01024; -;  
 CC DR Genew; HGNC:1318; C3.  
 CC MIM: 120700; -;  
 CC GO: GO:0005102; F:receptor binding activity; TAS.  
 CC GO: GO:0007186; P:G-protein coupled receptor protein signaln. . .; TAS.  
 CC DR GO: GO:0006955; P:Immune response; TAS.  
 CC DR GO: GO:0007165; P:signal transduction; TAS.  
 CC DR InterPro: IPR002890; A2M N.  
 CC DR InterPro: IPR000020; Anaphylatoxin.  
 CC DR InterPro: IPR001840; Anaphylatoxin.  
 CC DR InterPro: IPR001599; Macroglobulin2.  
 CC DR InterPro: IPR001134; Netrin\_C.  
 CC DR Pfam: PF00207; A2M; 1.  
 CC DR Pfam: PF01835; A2M N; 1.  
 CC DR Pfam: PF01821; ANATO; 1.  
 CC DR Pfam: PF01759; NTR; 1.  
 CC DR PRINTS: PR00004; ANAPHYLATOXN.  
 CC DR PRODOM: PD003264; Anaphylatoxin; 1.  
 CC DR SMART; SM00104; ANATO; 1.  
 CC DR SMART; SM00643; C345C; 1.  
 CC DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 CC DR PROSITE; PS01178; ANAPHYLATOXIN\_2; 1.  
 CC DR PROSITE; PS01177; ANAPHYLATOXIN\_1; 1.  
 CC Complement pathway; Complement alternate pathway; Plasma;  
 CC Inflammatory response; Glycoprotein; Signal; Polymorphism;  
 CC Disease mutation; 3D-structure; Thioester bond.  
 CC SIGNAL  
 CC 1 22  
 CC CHAIN 23 1663 COMPLEMENT C3.  
 CC CHAIN 23 667 COMPLEMENT C3, BETA CHAIN.  
 CC CHAIN 672 1663 COMPLEMENT C3, ALPHA CHAIN.  
 CC PEPTIDE 672 748 C3A ANAPHYLATOXIN.  
 CC CHAIN 749 1663 C3B ANAPHYLATOXIN.  
 CC PEPTIDE 749 954 C3C ALPHA' CHAIN.  
 CC PEPTIDE 749 954 C3C FRAGMENT.  
 CC PEPTIDE 955 1303 C3DG FRAGMENT.  
 CC PEPTIDE 955 1001 C3G FRAGMENT.  
 CC PEPTIDE 1002 1303 C3D FRAGMENT.  
 CC PEPTIDE 1304 1320 C3F FRAGMENT.  
 CC SITE 748 749 CLEAVAGE (BY C3 CONVERTASE).  
 CC SITE 954 955 CLEAVAGE (BY FACTOR I) (POTENTIAL).  
 CC SITE 1303 1304 CLEAVAGE (BY FACTOR I).  
 CC SITE 1320 1321 CLEAVAGE (BY FACTOR I).  
 CC DOMAIN 693 728 ANAPHYLATOXIN-LIKE.

Query Match 100.0%; Score 70; DB 1; Length 1663;  
 Best Local Similarity 100.0%; Pred. No. 0.00023;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITHRIHESASL 13  
 DB 1307 ITHRIHESASL 1319

RESULT 2  
 ID CO3 RABIT STANDARD; PRT; 726 AA.  
 AC P12247;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Complement C3 alpha chain (Fragment).  
 GN C3.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87006907; Pubmed=3019881;  
 RA Kueano M., Choi N.H., Tomita M., Yamamoto K., Migita S., Sekiya T.,  
 RA Nishimura S.;  
 RT "Nucleotide sequence of cDNA and derived amino acid sequence of  
 RT rabbit complement component C3 alpha-chain.";  
 RL Immunol. Invest. 15:365-378(1986).  
 CC -1- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE  
 CC COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL  
 CC REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS.  
 CC AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE  
 CC THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.  
 CC -1- SUBUNIT: C3 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 ARG  
 CC RESIDUES, FORMING TWO CHAINS, BETA & ALPHA, LINKED BY A DISULFIDE  
 CC BOND. C3 CONVERTASE ACTIVATES C3 BY CLEAVING THE ALPHA CHAIN,  
 CC RELEASING C3A ANAPHYLATOXIN & GENERATING C3B (BETA CHAIN + ALPHA  
 CC CHAIN).  
 CC -1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL, M32434; AAA31190.1; -;  
 DR PIR; A27602; A27602.  
 DR HSSP; P01024; 1C3D.  
 DR InterPro: IPR000020; Anaphylatoxin.  
 DR InterPro: IPR001599; Macroglobulin2.  
 DR InterPro: IPR001134; Netrin\_C.  
 DR Pfam; PF00207; A2M; 1.  
 DR Pfam; PF01759; NTR; 1.  
 DR SMART; SM00643; C34SC; 1.  
 DR PROSITE; PS01177; ANAPHYLATOXIN\_1; PARTIAL.  
 DR PROSITE; PS01178; ANAPHYLATOXIN\_2; PARTIAL.  
 DR PROSITE; PS00477; ALPHA\_2-MACROGLOBULIN; 1.  
 KW Complement pathway; Complement alternate pathway; Plasma;  
 KW Inflammatory response; Glycoprotein; Thioester bond.  
 FT NON TER 1  
 FT CHAIN 1  
 FT CROSSLINK <1 726  
 FT CARBOHYD 2 76  
 FT CARBOHYD 233 2  
 FT CARBOHYD 233 233  
 FT CARBOHYD 680 233  
 FT CARBOHYD 680 680  
 FT SEQUENCE 726 AA; 81844 MW; FAB4C35D461300B9 CRC64;

Query Match 70.0%; Score 49; DB 1; Length 726;  
 Best Local Similarity 69.2%; Pred. No. 0.48;  
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ITHRIHESASL 13  
 DB 370 VKHRIWDSASL 382

RESULT 3  
 ID SMP3\_YEAST STANDARD; PRT; 516 AA.  
 AC Q04174; Q99400;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE SMP3 protein.  
 GN SMP3 OR YOR149C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NBMS;  
 RX MEDLINE=91172125; Pubmed=2005867;  
 RA Irie K., Araki H., Oshima Y.;  
 RT "Mutations in a Saccharomyces cerevisiae host showing increased  
 RT holding stability of the heterologous plasmid psk1".  
 RL Mol. Gen. Genet. 225:257-265(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=6288C / FY1678;  
 RA Ayadi A., Bordome R., Camases A., Madania A., Poch O.,  
 RA Tarasov I.A., Winzor B., Martin R.P.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: ESSENTIAL PROTEIN INVOLVED IN PLASMID MAINTENANCE WITH  
 CC SMP2.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -----  
 CC -1- SIMILARITY: TO S.POMBE SPAC468.12C.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL, X58121; CAA41123.1; -;  
 DR EMBL; U55020; AAC49635.1; -;  
 DR EMBL; Z75057; CAA93355.1; -;  
 DR PIR; S67037; S67037.  
 DR SGP; S0005675; SMP3.  
 DR GO; GO:0006276; P:Plasmid maintenance; IMP.  
 DR InterPro: IPR005598; PMP.  
 DR Pfam; PF03901; PMP; 1.  
 KW Transmembrane.  
 FT TRANSMEM 6  
 FT TRANSMEM 26  
 FT TRANSMEM 61  
 FT TRANSMEM 176  
 FT TRANSMEM 211  
 FT TRANSMEM 271  
 FT TRANSMEM 291  
 FT TRANSMEM 296  
 FT TRANSMEM 318  
 FT TRANSMEM 349  
 FT TRANSMEM 369  
 FT TRANSMEM 122  
 FT TRANSMEM 123  
 FT TRANSMEM 163  
 FT TRANSMEM 169  
 FT TRANSMEM 279  
 FT TRANSMEM 516 AA; 59900 MW; 8D404622CB69534 CRC64;  
 FT CONFLICT 163 163  
 FT CONFLICT 169 169  
 FT CONFLICT 279 279  
 FT SEQUENCE 516 AA; 62.3%; Score 44; DB 1; Length 516;

Best Local Similarity 63.6%; Pred. No. 2.5;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 HRHMSASLL 13  
DB 207 YRHWKSPSL 217

## RESULT 4

RT09\_HUMAN STANDARD; PRT; 396 AA.  
AC P82933;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE 28S ribosomal protein S9, mitochondrial precursor (MRP-S9).  
GN MRP59 OR RPN59.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]

SEQUENCE FROM N.A.  
RA MEDLINE=22388257; PubMed=12477932;  
RA Strauberg R.L., Feigold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stempleton M., Soares W.B., Bonaldo W.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrincci P., Prange C.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Ravey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
RA Schermer A., Schin J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]

## IDENTIFICATION.

RA MEDLINE=21276436; PubMed=11279123;  
RA Koc E.C., Burkhardt W., Blackburn K., Moseley A., Spremulli L.L.;

RT "The small subunit of the mammalian mitochondrial ribosome:  
RT identification of the full complement of ribosomal proteins present.";

RL J. Biol. Chem. 276:19363-19374(2001).  
CC -1- SUBCELLULAR LOCATION: Mitochondrial.

CC -1- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.  
-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

CC EMBL; BF034318; -; NOT ANNOTATED CDS.  
DR InterPro; IPR000754; Ribosomal\_S9.

DR Pfam; PF00380; Ribosomal\_S9; 1.  
DR ProDom; PD001627; Ribosomal\_S9; 1.

DR PROSITE; PS00360; RIBOSOMAL\_S9; 1.  
KW Ribosomal protein; Mitochondrion; Transient peptide.

FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).  
FT CHAIN 1 ? 28S RIBOSOMAL PROTEIN S9.

FT SEQUENCE 396 AA; 45822 MW; A4ECC6DF37FE9AE CRC64;

Query Match 57.1%; Score 40; DB 1; Length 396;

Best Local Similarity 54.5%; Pred. No. 9.6;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 3 HRHMSASLL 13  
DB 175 HSHWQAKSL 185

## RESULT 5

PTXD\_PSEST STANDARD; PRT; 336 AA.  
AC O69054;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Phosphonate dehydrogenase (EC 1.20.1.1) (NAD-dependent phosphite  
DE dehydrogenase).  
GN PTXD.  
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=316;  
[1]

SEQUENCE FROM N.A.  
RA STRAIN=WM88;  
RA MEDLINE=99008986; PubMed=9791102;  
RA Metcalf W.W., Wolfe R.S.,  
RT "Molecular genetic analysis of phosphite and hypophosphite oxidation  
RT by Pseudomonas stutzeri WM88.";  
RL J. Bacteriol. 180:5547-5558(1998).  
[2]

SEQUENCE OF 1-15, FUNCTION, ACTIVITY, COFACTOR, ENZYME REGULATION,  
RP SUBUNIT, INDUCTION, AND MASS SPECTROMETRY.  
RC STRAIN=WM88;

RA MEDLINE=21264507; PubMed=11278981;  
RA Coates A.M.G., White A.K., Metcalf W.W.;

RT "Purification and characterization of a novel phosphorus-oxidizing  
RT enzyme from Pseudomonas stutzeri WM88.";

RL J. Biol. Chem. 276:17429-17436(2001).  
CC -1- FUNCTION: Catalyzes phosphite (phosphonate) oxidation.

CC -1- CATALYTIC ACTIVITY: Phosphonate + NAD(+) + H(2)O = phosphate +  
CC NADH.

CC -1- ENZYME REGULATION: Inhibited by NaCl, NADH and sulfite.  
CC -1- SUBUNIT: Homodimer.

CC -1- INDUCTION: By phosphate starvation.  
CC -1- MASS SPECTROMETRY: MW=36413; MW ERR=18; METHOD=MALDI.

CC -1- MISCELLANEOUS: Its optimum pH is between 7.25 and 7.75 and optimum  
CC temperature is 35 degrees Celsius.

CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID  
CC DEHYDROGENASES FAMILY.  
-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

CC EMBL; AF061070; AAC1709.1; -.  
DR HSPB; P36234; 1GDH.

DR InterPro; IPR006139; 2-Hacid DH.  
DR InterPro; IPR006140; 2-Hacid DH\_C.

DR Pfam; PF00389; 2-Hacid DH; 1.  
DR Pfam; PF02826; 2-Hacid DH\_C; 1.

DR PROSITE; PS00065; D\_2-HYDROXYACID\_DH\_1; FALSE NEG.  
DR PROSITE; PS00670; D\_2-HYDROXYACID\_DH\_2; FALSE NEG.

DR PROSITE; PS00671; D\_2-HYDROXYACID\_DH\_3; FALSE NEG.  
KW Oxidoreductase; NAD.

FT ACT\_SITE 237 237 SUBSTRATE BINDING (BY SIMILARITY).  
FT ACT\_SITE 266 266 BY SIMILARITY.

FT ACT\_SITE 292 292 BY SIMILARITY.  
FT SEQUENCE 336 AA; 36415 MW; 7F55D246CA4454F7 CRC64;



RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salberg S.L.,  
 RA Delcher A., Utechtack T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains." to the EMBL/GenBank/DBJ databases.  
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE UPF0073 (HLY-III) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AL021897; CAA17201.1; -;  
 CC EMBL; AE006992; AAK5373.1; -;  
 CC PIR; C70895; C70895.  
 CC TIGR; MT1117; -;  
 CC DR Tuberculat; Rv1085c; -;  
 CC DR InterPro; IPR004254; HlyIII-related.  
 CC DR InterPro; IPR005744; HlyIII.  
 CC DR Pfam; PF03006; UPF0073; 1.  
 CC DR TIGRFAMs; TIGR01065; hlyIII; 1.  
 CC KM Hypothetical protein; Transmembrane; Complete proteome.  
 CC FT TRANSMEM 42 62 POTENTIAL.  
 CC FT TRANSMEM 67 87 POTENTIAL.  
 CC FT TRANSMEM 108 128 POTENTIAL.  
 CC FT TRANSMEM 133 153 POTENTIAL.  
 CC FT TRANSMEM 159 179 POTENTIAL.  
 CC FT TRANSMEM 186 206 POTENTIAL.  
 CC FT TRANSMEM 222 242 POTENTIAL.  
 CC SQ SEQUENCE 242 AA; 26034 MW; BBC1DE12CF8D3500 CRC64;

Query Match 54.3%; Score 38; DB 1; Length 242;  
 Best Local Similarity 55.6%; Pred. No. 13;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 HRHMESAS 11  
 Db 89 HRVWMSKAT 97

RESULT 8  
 GEM2\_HUMAN STANDARD; PRT; 280 AA.  
 AC 014893;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Survival of motor neuron protein-interacting protein 1 (SMN-  
 DE interacting protein 1) (Component of gems 2) (Gemin2).  
 GN S1P1 OR GEMIN2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Euteleostomi; Primates; Catarrhini; Homidae; Homo.  
 CX NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast cancer;  
 RX MEDLINE=97462902; PubMed=93233129;  
 RA Liu Q., Fischer U., Wang F., Dreyfuss G.;  
 RT "The spinal muscular atrophy disease gene product, SMN, and its  
 RT associated protein S1P1 are in a complex with spliceosomal snRNP  
 RT proteins.";  
 RL Cell 90:1013-1021 (1997).  
 CC -1- FUNCTION: THE SMN COMPLEX PLAYS AN ESSENTIAL ROLE IN SPLICEOSOMAL  
 CC snRNP ASSEMBLY IN THE CYTOPLASM AND IS REQUIRED FOR PRE-mRNA  
 CC SPLICING IN THE NUCLEUS.  
 CC -1- SUBUNIT: FORMS A STABLE HETEROERIC COMPLEX WITH SURVIVAL OF MOTOR  
 CC NEURON PROTEIN (SMN), GEMIN3 AND GEMIN4.

CC -1- SUBCELLULAR LOCATION: LOCALIZED IN SUBNUCLEAR STRUCTURES NEXT TO  
 CC COILED BODIES, CALLED GEMS, WHICH ARE HIGHLY ENRICHED IN  
 CC SPLICEOSOMAL snRNPs. ALSO FOUND IN THE CYTOPLASM.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF027150; AAB82287.1; -;  
 CC Gene; HGNC:10884; S1P1.  
 CC DR GK; 014893; -;  
 CC DR MIM; 602595; -;  
 CC DR GO; GO:0005737; C:cytoplasm; TAS.  
 CC DR GO; GO:0005681; C:spliceosome complex; TAS.  
 CC DR GO; GO:0008248; F:pre-mRNA splicing factor activity; TAS.  
 CC DR GO; GO:0006397; P:mRNA processing; TAS.  
 CC DR GO; GO:0000245; P:spliceosome assembly; TAS.  
 CC DR InterPro; IPR007022; S1P1.  
 CC DR Pfam; PF04938; S1P1; 1.  
 CC KM mRNA processing; Spliceosome; Nuclear protein.  
 CC FT DOMAIN 101 106 POLY-GLN.  
 CC SQ SEQUENCE 280 AA; 31585 MW; 3232P410EA98EB81 CRC64;

Query Match 54.3%; Score 38; DB 1; Length 280;  
 Best Local Similarity 60.0%; Pred. No. 15;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 HRHMESASL 12  
 Db 120 HRSWMSQQL 129

RESULT 9  
 YHM7\_YEAST STANDARD; PRT; 280 AA.  
 ID YHM7\_YEAST  
 AC P38790;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Hypothetical 33.1 kDa protein in SSp1-DYS1 intergenic region.  
 GN YHR067W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 CX NCB1\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=94378003; PubMed=8091229;  
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,  
 RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,  
 RA Kucaba T., Hillier L., Jier M., Johnston L., Langston V.,  
 RA Lacroix P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,  
 RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaeth E., Vaughan K.,  
 RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,  
 RA Vaudin M.;  
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome  
 RT VIII.";  
 RL Science 265:2077-2082 (1994).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U00061; AAB6378.1; -;



DR PIR; S46699; S46699.  
 DR SGD; S0001109; YHR067W.  
 KM Hypothetical protein.  
 SQ SEQUENCE 280 AA; 33055 MW; 9FE7B9A602B7A083 CRC64;

Query Match 54.3%; Score 38; DB 1; Length 280;  
 Best Local Similarity 83.3%; Pred. No. 15;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 HRIHWE 8  
 DB 188 HRIHWD 193

RESULT 10  
 GLMS\_WIGBR STANDARD; PRT; 611 AA.  
 ID GLMS\_WIGBR  
 AC Q8D3D0;  
 DT 15-SEP-2003 (Rel. 42, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Glucosamine--fructose-6-phosphate aminotransferase (isomerizing)  
 DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-phosphate  
 DE phosphatase) (GFAV) (L-glutamine-D-fructose-6-phosphate  
 DE aminotransferase) (Glucosamine-6-phosphate synthase).  
 GN GLMS OR WIGBR0110.  
 OS *Wigglesworthia glossinidia brevipalpis*.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Wigglesworthia.  
 OX NCBI\_TaxID=36870;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22297718; PubMed=12219091;  
 RA Akman L., Yamashita A., Matanabe H., Oshima K., Shiba T., Hattori M.,  
 RA Aksoy S.;  
 RT "Genome sequence of the endocellular obligate symbiont of tsetse  
 RT flies, *Wigglesworthia glossinidia*.";  
 RL Nat. Genet. 32:402-407(2002).  
 CC -1- FUNCTION: Catalyzes the first step in hexosamine metabolism,  
 CC converting fructose-6P into glucosamine-6P using glutamine as a  
 CC nitrogen source (By similarity).  
 CC -1- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-  
 CC glutamate + D-glucosamine 6-phosphate.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SIS FAMILY.  
 CC GRAT SUBFAMILY.  
 CC -1- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR HMBP; AB063521; BAC24157.1; -  
 DR HMBP; MF\_00164; 1;  
 DR InterPro; IPR000583; GATase\_2.  
 DR InterPro; IPR005855; Gims.  
 DR InterPro; IPR001347; SIS.  
 DR Pfam; PF00310; GATase\_2; 1.  
 DR Pfam; PF01380; SIS\_2; 2;  
 DR TIGRfam; TIGR01135; gims; 1.  
 DR PROSITE; PS00443; GATASE\_TYPE\_II; 1.  
 DR Translase; Aminoacyltransferase; Glutamine amidotransferase;  
 KM Complete proteome.  
 FT INIT MET 0 BY SIMILARITY.  
 FT DOMAIN 1 243 GLUTAMINE AMIDOTRANSFERASE.  
 FT ACT\_SITE 1 1 GATASE (BY SIMILARITY).  
 FT ACT\_SITE 606 606 ISOMERIZATION FRU-6P (BY SIMILARITY).  
 SQ SEQUENCE 611 AA; 68790 MW; 26571D15AF705AAE CRC64;

Query Match 54.3%; Score 38; DB 1; Length 611;  
 Best Local Similarity 75.0%; Pred. No. 35;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ITHRIHWE 8  
 DB 129 IATHIHW 136

RESULT 11  
 DSCA\_HUMAN STANDARD; PRT; 2012 AA.  
 ID DSCA\_HUMAN  
 AC O60469; O60468;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Down syndrome cell adhesion molecule precursor (CHD2).  
 GN DSCAM.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC TISSUE=Brain;  
 RX MEDLINE=98087574; PubMed=9426258;  
 RA Yamakawa K., Huot Y.-K., Haendelt M.A., Hubert R., Chen X.-N.,  
 RA Lyons G.E., Korenberg U.R.;  
 RT "DSCAM: a novel member of the immunoglobulin superfamily maps in a  
 RT Down syndrome region and is involved in the development of the  
 RT nervous system.";  
 RL Hum. Mol. Genet. 7:227-237(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND FUNCTION.  
 RX MEDLINE=20384934; PubMed=10925149;  
 RA Agarwala K.L., Nakamura S., Tautsuni Y., Yamakawa K.;  
 RT "Down syndrome cell adhesion molecule DSCAM mediates homophilic  
 RT intercellular adhesion.";  
 RL Brain Res. Mol. Brain Res. 79:118-126(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20289799; PubMed=10830953;  
 RA Hattori M., Fujiyama A., Taylor T.D., Matanabe H., Yada T.,  
 RA Park H.-S., Toyoda A., Ishii K., Tokoki Y., Choi D.-K., Groner Y.,  
 RA Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,  
 RA Polley A., Menzel U., Delabar U., Kump K., Lehmann R., Paterson D.,  
 RA Reischwald K., Rump A., Schillabel M., Schudy A., Zimmermann W.,  
 RA Rosenthal A., Kudo T., Shibuya K., Kawasaki K., Asakawa S.,  
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,  
 RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,  
 RA Scharte M., Schoen O., Desario A., Reichelt J., Kauer G., Blocker H.,  
 RA Kanster U., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,  
 RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,  
 RA Lehrach H., Reinhardt R., Yaspo M.-L.;  
 RT "The DNA sequence of human chromosome 21.";  
 RL Nature 405:311-319(2000).  
 CC -1- FUNCTION: CELL ADHESION MOLECULE THAT CAN MEDIATE CATION-  
 CC INDEPENDENT HOMOPHILIC BINDING ACTIVITY. COULD BE INVOLVED IN  
 CC NERVOUS SYSTEM DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (PROBABLE). THE  
 CC SHORT ISOFORM MAY BE SECRETED.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Long; Synonyms=CHD2-52;  
 CC Name=Short; Synonyms=CHD2-42;  
 CC IsoId=O60469-2; Sequence=VSP\_002502, VSP\_002503;  
 CC -1- TISSUE SPECIFICITY: PRIMARILY EXPRESSED IN BRAIN.  
 CC -1- SIMILARITY: Contains 10 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 6 fibronectin type III domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC	the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.ebi.ac.uk/announcement/">http://www.ebi.ac.uk/announcement/</a> or send an email to <a href="mailto:license@ebi.ac.uk">license@ebi.ac.uk</a> ).
CC	-----
DR	EMBL; AF023450; AAC17967.1; -
DR	EMBL; AF023449; AAC17966.1; -
DR	EMBL; AF217525; AAF27525.1; -
DR	EMBL; AL163283; CAB90464.1; -
DR	EMBL; AL163282; CAB90436.1; -
DR	EMBL; AL163281; CAB90444.1; -
DR	Genew; HGNC:3039; DSCAM.
DR	MIM; 602523; -
DR	GO; GO:0005887; C:integral to plasma membrane; TAS.
DR	GO; GO:0005624; C:membrane fraction; TAS.
DR	GO; GO:0007155; P:cell adhesion; TAS.
DR	GO; GO:0007399; P:neurogenesis; TAS.
DR	InterPro; IPR003961; FN-III.
DR	InterPro; IPR007110; Ig-like.
DR	InterPro; IPR003598; Ig_c2.
DR	InterPro; IPR003006; Ig_MHC.
DR	Pfam; PF00047; Ig; 10.
DR	Pfam; PF00047; fn3; 6.
DR	SMART; SMO0060; FN3; 6.
DR	SMART; SMO0408; IGC2; 7.
DR	PROSITE; PS50835; IG_LIKE; 9.
KW	Immunoglobulin domain; Glycoprotein; signal; Cell adhesion; Repeat; Transmembrane; Alternative splicing.
FT	SIGNAL 1 17 POTENTIAL.
FT	CHAIN 18 2012 DOWN SYNDROME CELL ADHESION MOLECULE.
FT	DOMAIN 18 1595 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 1596 1616 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 1617 2012 POTENTIAL.
FT	DOMAIN 139 129 IG-LIKE C2-TYPE 1.
FT	DOMAIN 125 216 IG-LIKE C2-TYPE 2.
FT	DOMAIN 225 305 IG-LIKE C2-TYPE 3.
FT	DOMAIN 313 401 IG-LIKE C2-TYPE 4.
FT	DOMAIN 407 500 IG-LIKE C2-TYPE 5.
FT	DOMAIN 504 592 IG-LIKE C2-TYPE 6.
FT	DOMAIN 596 685 IG-LIKE C2-TYPE 7.
FT	DOMAIN 690 783 IG-LIKE C2-TYPE 8.
FT	DOMAIN 787 883 IG-LIKE C2-TYPE 9.
FT	DOMAIN 885 972 FIBONECTIN TYPE-III 1.
FT	DOMAIN 984 1076 FIBONECTIN TYPE-III 2.
FT	DOMAIN 1088 1177 FIBONECTIN TYPE-III 3.
FT	DOMAIN 1189 1273 FIBONECTIN TYPE-III 4.
FT	DOMAIN 1285 1377 IG-LIKE C2-TYPE 10.
FT	DOMAIN 1380 1463 FIBONECTIN TYPE-III 5.
FT	DOMAIN 1477 1562 FIBONECTIN TYPE-III 6.
FT	DISULFID 46 102 BY SIMILARITY.
FT	DISULFID 145 197 BY SIMILARITY.
FT	DISULFID 246 293 BY SIMILARITY.
FT	DISULFID 335 385 BY SIMILARITY.
FT	DISULFID 428 484 BY SIMILARITY.
FT	DISULFID 525 575 BY SIMILARITY.
FT	DISULFID 617 669 BY SIMILARITY.
FT	DISULFID 711 766 BY SIMILARITY.
FT	DISULFID 809 865 BY SIMILARITY.
FT	DISULFID 1307 1359 BY SIMILARITY.
FT	CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 470 470 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 512 512 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 556 556 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 658 658 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 666 666 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 710 710 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 748 748 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 795 795 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 924 924 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 1142 1142 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT	CARBOHYD	1160	1160	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	1250	1250	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	1271	1271	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	1341	1341	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	1488	1488	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	VARSPPLIC	1562	1571	NEFALTNDGS -> KEAARCKEFS (In isoform Short)	
FT				/PrtId=VSP_002502.	
FT	VARSPPLIC	1572	2012	Missing (In isoform Short).	
FT				/PrtId=VSP_002503.	
FT	CONFLICT	1693	2012	HRPDLIHLPPYLMDFLNRGPGPTGRDLSIGACLEPQK	
FT				SRTLKRPVLVEIPIMEASASSTRREGSQMGPVATLPOR	
FT				EGARLGOAATKMSQSESLDGRGLKKNPNPAKSYTLV ->	
FT				IGQVTSYICHTHTEWTRC (IN REF. 1).	
FT					
FT	SEQUENCE	2012 AA;	222259 MW;	0E33CFB781A08334 CRC64;	
Qy					
Db					
	1702 VTHRTVHYQSVS	1712			
RESULT 12					
GLMS_YERPE					
ID	GLMS_YERPE	STANDARD;	PRT;	608 AA.	
AC	082958;				
DT	28-FEB-2003 (Rel. 41, Created)				
DT	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Glucosamine-6-phosphate aminotransferase [isomerizing]				
DE	(EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-phosphate aminotransferase) (GFAF) (L-glutamine-D-fructose-6-phosphate				
DE	aminotransferase) (Glucosamine-6-phosphate synthase).				
GN	GLMS OR YPO4118 OR Y4132.				
OS	Yersinia pestis.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Yersinia.				
OX	NCBI_TaxId=632;				
RN	[1]				
RN	SEQUENCE FROM N.A.				
RP	STRAIN=CO-92 / Biovar Orientalis;				
RX	MEDLINE=21470413; PubMed=1156360;				
RA	Parkhill J., Wren B.W., Thomson N.R.,				
RA	Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,				
RA	Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,				
RA	Chillingworth T., Cronin A., Davies R.W., Davis P., Dougan G.,				
RA	Fellwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,				
RA	Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,				
RA	Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;				
RT	"Genome sequence of Yersinia pestis, the causative agent of plague."				
RL	Nature 413:523-527(2001).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=KIMS / Biovar Mediaevalis;				
RX	MEDLINE=22137863; PubMed=12142430;				
RA	Deng W., Burland V., Plunkett G., IIT, Boutin A., Mayhew G.F., Liss P.,				
RA	Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,				
RA	Feherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,				
RA	Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,				
RA	Perry R.D.;				
RT	"Genome sequence of Yersinia pestis KIM."				
RL	J. Bacteriol. 194:4601-4611(2002).				
CC	-1- FUNCTION: Catalyzes the first step in hexosamine metabolism,				
CC	converting fructose-6P into glucosamine-6P using glutamine as a				
CC	nitrogen source (By similarity).				
CC	-1- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-				
CC	glutamate + D-glucosamine 6-phosphate.				
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).				
CC	-1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE SIS FAMILY.				
CC	GPAI SUBFAMILY.				

CC -1- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AJ141416; CAC93567.1; -  
 DR EMBL; AE014014; AAM87674.1; -  
 DR PIR; AB0500; AB0500.  
 DR HAMAP; MF\_00164; -; 1.  
 DR InterPro; IPR000583; GATase\_2.  
 DR InterPro; IPR005855; GIMS.  
 DR InterPro; IPR01347; SIS.  
 DR Pfam; PF00310; GATase\_2; 1.  
 DR Pfam; PF01380; SIS; 2.  
 DR TIGRFAMs; TIGR01135; gims; 1.  
 DR PROSITE; PS00443; GATASE\_TYPE\_II; 1.  
 DR Transferrase; Aminotransferase; Glutamine amidotransferase;  
 KW Complete proteome.  
 FT INT MET 0 BY SIMILARITY.  
 FT DOMAIN 1 240 GLUTAMINE AMIDOTRANSFERASE.  
 FT ACT\_SITE 1 1 GATASE (BY SIMILARITY).  
 FT ACT\_SITE 603 603 ISOMERIZATION FRU-6P (BY SIMILARITY).  
 SQ SEQUENCE 608 AA; 66394 MW; 165FDDAD16CCD6 CRC64;  
 Query Match 53.9%; Score 37.5; DB 1; Length 608;  
 Best Local Similarity 50.0%; Pred. No. 42;  
 Matches 8; Conservative 1; Mismatches 4; Indels 3; Gaps 1;  
 Qy 1 ITHRIHME---SASLL 13  
 Db 127 IAHVHWEEQGGGSL 142  
 RESULT 13  
 PRT1\_PICAN STANDARD; PRT; 220 AA.  
 AC P12806;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 26-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative PRT1 protein.  
 GN PRT1.  
 OS Pichia angusta (Yeast) (Hansenula polymorpha).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Pichia.  
 OX NCBI\_TaxID=4905;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS 4732;  
 RX MEDLINE=89287321; PubMed=2500147;  
 RA Bruinenberg P.G., Evers M., Waterham H.R., Kuipers J., Arnborg A.C.,  
 Ab G.;  
 RT "Cloning and sequencing of the peroxisomal amine oxidase gene from  
 RT Hansenula polymorpha.";  
 RL Biochim. Biophys. Acta 1008:157-167(1989).  
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; X15111; CAA33208.1; -  
 DR PIR; S16511; S16511.  
 DR InterPro; IPR000504; RNA\_rec\_mot.

DR Pfam; PF00076; rrm; 1.  
 DR SMART; SM00360; RRM; 1.  
 DR PROSITE; PS0102; RRM; 1.  
 DR PROSITE; PS00030; RRM\_RNP\_1; FALSE\_NEG.  
 FT DOMAIN 37 120 RNA-BINDING (RRM).  
 SQ SEQUENCE 220 AA; 24961 MW; D317E7EFF49834B5 CRC64;  
 Query Match 53.9%; Score 37; DB 1; Length 220;  
 Best Local Similarity 58.3%; Pred. No. 17;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 Qy 1 ITHRIHMEASLL 12  
 Db 184 IEPRAHTSASM 195  
 RESULT 14  
 GEM2\_RAT STANDARD; PRT; 269 AA.  
 AC Q90ZP1;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Survival of motor neuron protein-interacting protein 1 (SMN-  
 DE interacting protein 1) (Component of gems 2) (Gemin2).  
 GN S1P1 OR GEMIN2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley;  
 RA Kallenbach S., La Bella V., Cisterni C., Bettmann B.;  
 RT "SMN interacting protein-1 expression pattern in rat.";  
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: THE SMN COMPLEX PLAYS AN ESSENTIAL ROLE IN SPliceosomal  
 CC SPLICING IN THE CYTOPLASM, AND IS REQUIRED FOR PRE-MRNA  
 CC SPLICING IN THE NUCLEUS (BY SIMILARITY).  
 CC -1- SUBUNIT: FORMS A STABLE HETEROMERIC COMPLEX WITH SURVIVAL OF MOTOR  
 CC NEURON PROTEIN (SMN), GEMIN3 AND GEMIN4 (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: LOCALIZED IN SUBNUCLEAR STRUCTURES NEXT TO  
 CC COILED BODIES, CALLED GEMS, WHICH ARE HIGHLY ENRICHED IN  
 CC SPliceosomal SNRNPs. ALSO FOUND IN THE CYTOPLASM (BY SIMILARITY).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AF176072; AAD53287.1; -  
 DR InterPro; IPR007022; S1P1.  
 DR Pfam; PF04938; S1P1; 1.  
 KW mRNA processing; Spliceosome; Nuclear protein.  
 FT DOMAIN 90 95 POLY-GLN.  
 SQ SEQUENCE 269 AA; 30440 MW; E03AB9DBFCC7EE3 CRC64;  
 Query Match 53.9%; Score 37; DB 1; Length 269;  
 Best Local Similarity 60.0%; Pred. No. 21;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 Qy 3 HRIHMEASLL 12  
 Db 109 HRNMKSOQL 118  
 RESULT 15  
 M1AA\_MYCTU STANDARD; PRT; 314 AA.



FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 587 AA; 67544 MW; 22FDAF8588FEC2AE CRC64;  
 Query Match 52.9%; Score 37; DB 1; Length 587;  
 Best Local Similarity 85.7%; Pred. No. 50;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 HRIHMS 9  
 Db 217 HRIHMS 223  
 RESULT 17  
 T9S3 HUMAN STANDARD; PRT: 589 AA.  
 AC Q9HDS; Q9NMH8; Q9P0G9; Q9UHM6;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transmembrane 9 superfamily protein member 3 precursor (SM-11044  
 binding protein) (EP70-P-Iso).  
 GN TM9SF3 OR SMP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Warner S.J., Lomax M.I.;  
 RT "Evolution of the TM9 super family of membrane spanning proteins.";  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Adrenal gland;  
 RA Gu Y., Peng Y., Li Y., Fu S., Gu J., Gu W., Jiang C., Yu Y., Han Z.,  
 RA Wang Y., Chen Z., Fu G.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 12-589 FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RA Sugawara T., Lenzen G., Simon S., Hidaka J., Cahen A., Guillaume J.L.,  
 RA Camoin L., Nahmias C., Strosberg A.D.;  
 RT "The iodocyanophanol and SM-11044 binding protein (SMP) belongs to  
 the emerging family of MP70 multispinning membrane proteins.";  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 128-589 FROM N.A.  
 RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,  
 RA Nakajima Y., Mitano T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,  
 RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE NONASPANIN (TM9SF) FAMILY.  
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FRAMESHIFTS  
 IN POSITIONS 184; 198; 370 AND 446.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC  
 CC EMBL; AF69150; AAF69159.1; -  
 CC EMBL; AF160213; AAF67014.1; ALT\_FRAME.  
 CC EMBL; AF16347; AAF21983.1;  
 CC EMBL; AK00756; BAA91362.1; ALT\_INIT.  
 CC InterPro; IPR004240; EMP70.  
 CC Pfam; PF02990; EMP70.1.  
 CC Signal; Transmembrane.

FT SIGNAL 1 28 POTENTIAL.  
 FT CHAIN 29 589 TRANSMEMBRANE 9 SUPERFAMILY PROTEIN  
 FT TRANSMEM 224 244 POTENTIAL.  
 FT TRANSMEM 294 314 POTENTIAL.  
 FT TRANSMEM 328 348 POTENTIAL.  
 FT TRANSMEM 360 380 POTENTIAL.  
 FT TRANSMEM 389 409 POTENTIAL.  
 FT TRANSMEM 449 469 POTENTIAL.  
 FT TRANSMEM 482 502 POTENTIAL.  
 FT TRANSMEM 519 539 POTENTIAL.  
 FT TRANSMEM 551 571 POTENTIAL.  
 FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 15 15 MISSING (IN REF. 2).  
 FT CONFLICT 191 191 M->I (IN REF. 1).  
 FT CONFLICT 210 218 KYDPSFQ->NIIIVFS (IN REF. 2).  
 FT CONFLICT 232 232 M->T (IN REF. 4).  
 FT CONFLICT 547 547 F->S (IN REF. 4).  
 SQ SEQUENCE 589 AA; 67888 MW; E673D547A35BCA47 CRC64;  
 Query Match 52.9%; Score 37; DB 1; Length 589;  
 Best Local Similarity 85.7%; Pred. No. 50;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 HRIHMS 9  
 Db 219 HRIHMS 225  
 RESULT 18  
 TNPX\_ECOLI STANDARD; PRT: 698 AA.  
 ID Q00042;  
 AC Q00042;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Putative transposon gamma-delta 80.3 kDa protein (Transposon Tn1000  
 tnpX protein).  
 GN TNPX.  
 OS Escherichia coli.  
 OC Bacteriia; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95337425; PubMed=7612932;  
 RA Broom J.E., Hill D.F., Hughes G., Jones W.A., McNaughton J.C.,  
 RA Stockwell P.A., Petersen G.B.;  
 RT "Sequence of a transposon identified as Tn1000 (gamma delta).";  
 RL DNA Seq. 5:185-189(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94361836; PubMed=8080658;  
 RA Maekawa T., Ohsudo E.;  
 RT "Identification of the region that determines the specificity of  
 binding of the transposases encoded by Tn3 and gamma delta to the  
 terminal inverted repeat sequences.";  
 RL Jpn. J. Genet. 69:269-285(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / CR63;  
 RA Shimizu H., Saitoh Y., Suda Y., Uehara K., Sempel G., Mizobuchi K.;  
 RT "Complete nucleotide sequence of the P plasmid: its implications for  
 organization and diversification of plasmid genomes.";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; X60200; CAA42758.1; -.  
DR EMBL; D16449; BAA03916.1; -.  
DR EMBL; AP001918; BAA97878.1; -.  
DR PIR; I60218; I60218.  
DR EcGene; EG40023; tnpX.  
KW Hypothetical protein; Transposable element; Plasmid;  
KW Complete proteome.  
SQ SEQUENCE 698 AA; 80329 MW; 91826FB4F5575BB9 CRC64;

Query Match 52.9%; Score 37; DB 1; Length 698;  
Best Local Similarity 60.0%; Pred. No. 60;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 RHMESASL 13  
|:|:|:|:  
267 RLHWDMAQL 276

Db 267 RLHWDMAQL 276

RESULT 19  
OBP\_HSV11 STANDARD; PRT; 851 AA.  
ID OBP\_HSV11  
AC P10193;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Origin of replication binding protein.  
GN UL9.  
OS Herpes simplex virus (type 1 / strain 17).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
NC NCB1\_TaxID=10299;  
LN [1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=88274327; PubMed=2839594;  
RX McGeoch D.J., Dairymple M.A., Davison A.J., Dolan A., Frame M.C.,  
RA McComb D., Perry L.J., Scott J.E., Taylor P.;  
RT "The complete DNA sequence of the long unique region in the genome of  
RT herpes simplex virus type 1.";  
RL J. Gen. Virol. 69:1531-1574(1988).  
LN [2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=88091053; PubMed=2826807;  
RX McGeoch D.J., Dairymple M.A., Dolan A., McComb D., Perry L.J.,  
RA Taylor P., Challberg M.D.;  
RT "Structures of herpes simplex virus type 1 genes required for  
RT replication of virus DNA.";  
RL J. Virol. 62:444-453(1988).  
CC -!- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION. BINDS THE ORIGIN  
CC OF REPLICATION (ORI).  
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL9,  
CC EBV-1 53, AND VZV 51.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; X14112; CAA32345.1; -.  
DR EMBL; M19120; AAA45822.1; -.  
DR PIR; B29890; WMBEU9.  
DR TRANSFAC; T00957; -.  
DR InterPro; IPR001410; DEAD.  
DR InterPro; IPR003450; Herpes\_ori\_bp.  
DR Pfam; PF02399; Herpes\_ori\_bp; 1.  
DR SMART; SM00487; DEXDC\_1.  
KW DNA replication; DNA-binding; ATP-binding.  
NP\_BIND 81 88 ATP (POTENTIAL).

SQ SEQUENCE 851 AA; 94261 MW; 961A13FE7A30CA7 CRC64;

Query Match 52.9%; Score 37; DB 1; Length 851;  
Best Local Similarity 62.5%; Pred. No. 75;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 RHMESAS 11  
|:|:|:|:  
653 RLHWDMA 660

Db 653 RLHWDMA 660

RESULT 20  
AT14\_HUMAN STANDARD; PRT; 1223 AA.  
ID AT14\_HUMAN  
AC O8WXS8; O8TE55; O8TEY8;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE ADAMTS-14 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
DE with thrombospondin motifs 14) (ADAM-TS 14) (ADAM-TS14).  
GN ADAMTS14.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NC NCB1\_TaxID=9606;  
LN [1]  
RN SEQUENCE FROM N.A. (ISOFORM A).  
RP MEDLINE=21638061; PubMed=11779638;  
RX Bolz H., Ramirez A., von Brederlow B., Kubisch C.;  
RT "Characterization of ADAMTS14, a novel member of the ADAMTS  
RT metalloproteinase family.";  
RL Biochim. Biophys. Acta 1522:221-225(2001).  
LN [2]  
RN SEQUENCE FROM N.A. (ISOFORM A).  
RP TISSUE=Fetal lung;  
RX MEDLINE=21856482; PubMed=11867212;  
RX Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,  
RA Lopez-Otin C.;  
RT "Cloning, expression analysis, and structural characterization of  
RT seven novel human ADAMTS, a family of metalloproteinases with  
RT disintegrin and thrombospondin-1 domains.";  
RL Gene 283:49-62(2002).  
LN [3]  
RN SEQUENCE OF 29-1223 FROM N.A. (ISOFORMS B; C AND D), AND  
RP ALTERNATIVE PROMOTER USAGE.  
RX MEDLINE=21839041; PubMed=11741898;  
RA Collige A., Vandenberghe I., Thiry M., Lambert C.A., Van Beeumen J.,  
RA Li S.-W., Prockop D.J., Lapierre C.M., Nussgens B.V.;  
RT "Cloning and characterization of ADAMTS-14, a novel ADAMTS displaying  
RT high homology with ADAMTS-2 and ADAMTS-3.";  
RL J. Biol. Chem. 277:5756-5766(2002).  
CC -!- FUNCTION: Has a aminoproteoglycan type I activity processing  
CC activity in the absence of ADAMTS2. Seems to be synthesized as a  
CC latent enzyme that requires activation to display aminoproteoglycan  
CC peptidase activity.  
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
CC matrix (by similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative promoter;  
CC Comment=2 isoforms, A (shown here) and B, are produced by use of  
CC alternative promoters;  
CC Event=Alternative splicing; Named isoforms=4;  
CC Name=A;  
CC IsoId=Q8WXS8-1; Sequence=Displayed;  
CC Name=B;  
CC IsoId=Q8WXS8-2; Sequence=VSP\_006958;  
CC Name=C;  
CC IsoId=Q8WXS8-3; Sequence=VSP\_006958, VSP\_005501;  
CC Note=Produced by alternative splicing of isoform B;  
CC Name=D;  
CC IsoId=Q8WXS8-4; Sequence=VSP\_005501;  
CC Note=Produced by alternative splicing of isoform A;  
CC -!- TISSUE SPECIFICITY: Expressed in retina and at low levels in

CC brain, lung and placenta. High expression in fetal tissues.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important  
 CC for a tight interaction with the extracellular matrix (By  
 CC similarity).  
 CC -1- PFM: The precursor is cleaved by a furin endopeptidase (By  
 CC similarity).  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 diintegrin-like domain.  
 CC -1- SIMILARITY: Contains 1 PLAC domain.  
 CC -1- SIMILARITY: Contains 4 TSP type-1 domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF358666; AAL40229.1; -  
 CC EMBL: AJ345098; CAC87943.1; -  
 CC EMBL: AF366351; AAL79814.1; -  
 CC MEROPS: M12.024; -  
 CC Genew: HGNC:14899; ADAMTS14.  
 CC MIM: 607506; -  
 CC InterPro: IPR002870; Pep\_M12B\_propep.  
 CC InterPro: IPR001590; Reptolysin.  
 CC InterPro: IPR000884; TSP1.  
 CC Pfam: PF01562; Pep\_M12B\_propep; 1.  
 CC Pfam: PF01421; Reptolysin; 1.  
 CC Pfam: PF00090; tsp\_1; 4.  
 CC SMART: SMO0209; TSP1; 4.  
 CC PROSITE: PS50215; ADAM\_MEROPS; 1.  
 CC PROSITE: PS00546; CYSTEINE\_SWITCH; FALSE\_NEG.  
 CC PROSITE: PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 CC PROSITE: PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
 CC PROSITE: PS50092; TSP1; 4.  
 CC PROSITE: PS00142; ZINC\_PROTEASE; FALSE\_NEG.  
 CC Hydrolyase: Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 CC Collagen degradation; Repeat; Extracellular matrix;  
 CC Alternative splicing; Alternative promoter usage.  
 CC SIGNAL 1 22  
 CC PROPEP 23 252  
 CC CHAIN 253 1223  
 CC DOMAIN 253 460  
 CC DOMAIN 461 551  
 CC DOMAIN 552 607  
 CC DOMAIN 730 846  
 CC DOMAIN 847 907  
 CC DOMAIN 908 967  
 CC DOMAIN 968 1022  
 CC DOMAIN 1063 1093  
 CC DOMAIN 729 739  
 CC DOMAIN 608 729  
 CC DOMAIN 875 878  
 CC DOMAIN 1100 1221  
 CC ACT\_SITE 399 399  
 CC METAL 398 398  
 CC METAL 402 402  
 CC METAL 408 408  
 CC CARBOHYD 109 109  
 CC CARBOHYD 475 475  
 CC CARBOHYD 941 941  
 CC CARBOHYD 1027 1027  
 CC VARSPIC 1 68  
 CC FT 368 368  
 CC FT 868 868  
 CC FT 884 884  
 CC FT 901 901  
 CC FT 923 923  
 CC FT 1024 1024  
 CC FT 1223 AA; 133871 MW; 36394AC4D92F170F CRC64;  
 CC SEQUENCE

Query Match 52.9%; Score 37; DB 1; Length 1223;  
 Best Local Similarity 54.5%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 Qy 3 HRHMSASL 13  
 Db 434 HRHMSASL 444  
 RESULT 21  
 ID RMP2\_HUMAN STANDARD; PRT; 175 AA.  
 AC O60895; OSN1P2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Receptor activity-modifying protein 2 precursor (CR1R activity-  
 DE modifying-protein 2) (Calcitonin-receptor-like receptor-activity-  
 DE modifying-protein 2).  
 GN RAMP2  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.A. FUNCTION, AND TISSUE SPECIFICITY.  
 RC TISSUE=Neuroblastoma;  
 RX MEDLINE=98282119; PubMed=9620797;  
 RA McLatchie L.M., Fraser N.J., Main M.J., Wise A., Brown J.,  
 RA Thompson N., Solari R., Lee M.G., Foord S.M.;  
 RT "RAMPs regulate the transport and ligand specificity of the  
 RL calcitonin-receptor-like receptor.";  
 RL Nature 393:333-339(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.A.  
 RC TISSUE=Pancreas;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,  
 RA Diatchenko L., Marisela K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Transports calcitonin-receptor-like receptor (CR1R) to  
 CC the plasma membrane where it acts as an adrenergic receptor.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN LUNG, BREAST, IMMUNE  
 CC SYSTEM AND FETAL TISSUES.  
 CC -1- SIMILARITY: BELONGS TO THE RAMP FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AJ001015; CAA0473.1; -  
 DR EMBL; BC027975; AAH27975.1; -  
 DR Genew; HGNC:9844; RAMP2.  
 DR MIM; 605154; -  
 DR GO; GO:0005905; C:coated pit; TAS.  
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
 DR GO; GO:0005764; C:lysosome; TAS.  
 DR GO; GO:0015031; P:protein transport; TAS.  
 DR GO; GO:0006898; P:receptor mediated endocytosis; TAS.  
 DR InterPro; IPR006985; RAMP.  
 DR Pfam; PF04901; RAMP; 1.  
 DR Signal; Transmembrane; Transport; Receptor.  
 DR SIGNAL; 1 42 POTENTIAL.  
 FT CHAIN 43 175 RECEPTOR ACTIVITY-MODIFYING PROTEIN 2.  
 FT DOMAIN 43 145 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 146 166 POTENTIAL.  
 FT DOMAIN 167 175 CYTOPLASMIC (POTENTIAL).  
 FT CONFLICT 25 25 L->V (IN REF. 1).  
 FT SEQUENCE 175 AA; 19607 MW; AFE9A9A461EFC6A3 CRC64;  
 Query Match 51.4%; Score 36; DB 1; Length 175;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 Qy 2 THRIHESASLL 13  
 Db 123 THQIHFNCSLV 134  
 RESULT 22  
 CAH1 MOUSE STANDARD; PRT; 260 AA.  
 ID CAH1\_MOUSE  
 AC P13634; Q9DC84;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Carbonic anhydrase I (EC 4.2.1.1) (Carbonate dehydratase I) (CA-I).  
 GN CAI OR CAR1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87169766; PubMed=3104601;  
 RA Fraser P.J., Curtis P.J.;  
 RT "Molecular evolution of the carbonic anhydrase genes: calculation of  
 RT divergence time for mouse carbonic anhydrase I and II.";  
 RL J. Mol. Evol. 23:294-299 (1986).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=90014784; PubMed=2571923;  
 RA Fraser P.J., Cummings P., Curtis P.J.;  
 RT "The mouse carbonic anhydrase I gene contains two tissue-specific  
 RT promoters.";  
 RL Mol. Cell. Biol. 9:3308-3313 (1989).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Spleen;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arkawa T., Hara A., Fukunishi Y., Kono H., Aachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner T., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schmitt L.M., Stuhli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaeys P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilting L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690 (2001).  
 [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colon;  
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.R.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshitsuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Muliyil S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalins D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -1- FUNCTION: Reversible hydration of carbon dioxide.  
 CC -1- CATALYTIC ACTIVITY: H(2)O + CO(2) = CO(3) + H(2)O.  
 CC -1- COFACTOR: Zinc.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase  
 CC family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL; M32452; AAA37354.1; -  
 DR EMBL; U36655; AAA50291.1; JOINED.  
 DR EMBL; M28197; AAA50291.1; JOINED.  
 DR EMBL; U36650; AAA50291.1; JOINED.  
 DR EMBL; U36651; AAA50291.1; JOINED.  
 DR EMBL; U36652; AAA50291.1; JOINED.  
 DR EMBL; U36653; AAA50291.1; JOINED.  
 DR EMBL; U36654; AAA50291.1; JOINED.  
 DR EMBL; AK003066; BAAB2544.1; -  
 DR EMBL; BC011223; AAH1223.1; -  
 DR PIR; A26344; A26344.  
 DR HSSP; P00915; IBLZM.  
 DR MGP; MGI:88258; Car1.  
 DR InterPro; IPR001148; Euk Coahnd.  
 DR Pfam; PF00194; carb\_anhydrase; 1.  
 DR PROSITE; PS000865; Euk Coahnd; 1.  
 DR PROSITE; PS00162; Euk CO2 ANHYDRASE; 1.  
 KW Lyase; Zinc; Metal-binding.  
 FT INIT\_MET 0  
 FT METAL 94  
 FT METAL 96  
 FT METAL 119  
 FT CONFLICT 237  
 FT SEQUENCE 260 AA; 28189 MW; AOC29A7BBBCFEFC CRC64;  
 Query Match 51.4%; Score 36; DB 1; Length 260;  
 Best Local Similarity 62.5%; Pred. No. 31;



Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 HRHMSA 10  
|:|:|:|  
Db 119 HLVMNSA 126

RESULT 23  
CAH1\_SHEEP STANDARD; PRT; 260 AA.  
AC P46282;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Carbonic anhydrase I (EC 4.2.1.1) (Carbonate dehydratase I) (CA-I).  
GN CAI.  
OS Ovis aries (sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Dorset; TISSUE=Ruminal epithelium;  
RX MEDLINE=96251459; PubMed=8690670;  
RA Wang L.O., Baldwin R.L., Jesse B.W.;  
RT "Isolation and characterization of a cDNA clone encoding ovine type I  
RT carbonic anhydrase.";  
RL J. Anim. Sci. 74:345-353 (1996).  
CC -1- FUNCTION: Reversible hydration of carbon dioxide.  
CC -1- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.  
CC -1- COFACTOR: zinc.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase  
CC family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: L42178; AAC1634.1; -.  
DR HSSP; P00915; 182M.  
DR InterPro; IPR001148; Euk\_COand.  
DR Pfam; PF00194; carb\_anhydrase; 1.  
DR ProDom; PD000865; Euk\_COand; 1.  
DR PROSITE; PS00162; EUK\_CO2\_ANHYDRASE; 1.  
KM Lyase; Zinc; Metal-binding.  
FT INIT\_MET 0 BY SIMILARITY.  
FT METAL 94 ZINC (CATALYTIC).  
FT METAL 96 ZINC (CATALYTIC).  
FT METAL 119 ZINC (CATALYTIC).  
SQ SEQUENCE 260 AA; 28766 MW; CB96A08F5CB041CE CRC64;

Query Match 51.4%; Score 36; DB 1; Length 260;  
Best Local Similarity 62.5%; Pred. No. 31;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 HRHMSA 10  
|:|:|:|  
Db 119 HLVMNSA 126

RESULT 24  
NOD1\_AZOCA STANDARD; PRT; 320 AA.  
AC Q07756;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Nodulation ATP-binding protein I.  
GN NOD1.  
OS Azorhizobium caulinodans.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Hyphomicrobiaceae; Azorhizobium.  
OX NCBI\_TaxID=7;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ORS571;  
RX MEDLINE=94018601; PubMed=8412659;  
RA Geelen D., Mergaert P., Geremta R.A., Goormachtig S., Montagu M.,  
RA Holsters M.;  
RT "Identification of nodSUT genes in Nod locus 1 of Azorhizobium  
RT caulinodans: evidence that nodS encodes a methyltransferase involved  
RT in Nod factor modification.";  
RL Mol. Microbiol. 9:145-154 (1993).  
CC -1- FUNCTION: FORMS, WITH NODU, A MEMBRANE TRANSPORT COMPLEX INVOLVED  
CC IN THE NODULATION PROCESS. IT PROBABLY EXPORT A MODIFIED BETA-1,4-  
CC LINKED N-ACETYLGLUCOSAMINE OLIGOSACCHARIDE.  
CC -1- SUBCELLULAR LOCATION: Membrane-associated.  
CC -1- SIMILARITY: Belongs to the ABC transporter family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: L18897; AAB51167.1; -.  
DR PIR; S35007; S35007.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR003439; ABC\_transporter.  
DR Pfam; PF00005; ABC\_tran; 1.  
DR ProDom; PD000006; ABC\_transporter; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; FALSE\_NEG.  
DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
KM Nodulation; Membrane; ATP-binding; Transport.  
FT NP\_BIND 47 ATP (BY SIMILARITY).  
FT NP\_BIND 54 ATP (BY SIMILARITY).  
SQ SEQUENCE 320 AA; 35310 MW; D89B3BF1E7CC4392 CRC64;

Query Match 51.4%; Score 36; DB 1; Length 320;  
Best Local Similarity 56.3%; Pred. No. 39;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 THRHMSASL 13  
|:|:|:|  
Db 222 TRAEWETALL 233

RESULT 25  
SEPT3\_HUMAN STANDARD; PRT; 345 AA.  
AC Q9UHO3;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Neuronal-specific septin 3.  
GN SEPT3 OR SEPT3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).  
RA Methner A., Leuwerenz J., Leyboldt F.;  
RT "Identification and characterization of human septin 3 on chromosome  
RT 22q13.2 upregulated by retinoic acid induced differentiation of the  
RT human neuronal precursor cell line Ntera/D2.";  
RL submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
[2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=20057165; PubMed=10591208;  
 RA Dunham I., Hunt A.R., Collins J.E., Brunkewich R., Beare D.M.,  
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,  
 RA Baguley C., Bailey U., Barlow K.F., Bates K.N., Beasley O.P.,  
 RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,  
 RA Burrill W.D., Burton J., Carter C., Carter N.P., Chen Y., Clark G.,  
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,  
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,  
 RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,  
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,  
 RA Gilbert J.G.R., Goward M.E., Graham D.V., Griffiths M.N.D., Hall C.,  
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,  
 RA Hunt S.E., Jones M.C., Kerhaw J., Kimberley A.M., King A.,  
 RA Laird G.K., Langford C.F., Leverisha M.A., Lloyd C., Lloyd D.M.,  
 RA Martyn I.D., Maehregh-Mohammadi M., Matthews L.H., Mccann O.T.,  
 RA McElay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.U.,  
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,  
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Rose M.T.,  
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,  
 RA Soderlund C., Spragon L., Steward C.A., Suleton J.E., Swann R.M.,  
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,  
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilting L.,  
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,  
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,  
 RA Roe B.A., Chen F., Chu Y., Crabtree J., Deschamps S., Do A., Do T.,  
 RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,  
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malai E., Nguyen T., Pan H.,  
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,  
 RA Phan Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,  
 RA Zhang M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,  
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,  
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,  
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozerik P., Rohlfing T.,  
 RA Schaefer P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,  
 RA Koef I., Bedell J.A., Hillier A., Mardis E., Waterston R., Wilson R.,  
 RA Emanuel H.E., Shaikh T., Kurahashi H., Saito S., Budarf M.L.,  
 RA McErmid H.E., Johnson A., Wong A.C.G., Morrow B.E., Edelmann L.,  
 RA Kim U.J., Shizuya H., Simon M.I., Dunneki J.P., Peyraud K.P., Ketra D.,  
 RA Seroussi E., Franssen I., Tapia I., Bruder C.E., O'Brien K.P.,  
 RA Wilkinson P., Bodenreich A., Hartman K., Hu X., Khan A.S., Lane L.,  
 RA Tiliakou Y., Wright H.,  
 RA "The DNA sequence of human chromosome 22.";  
 RT Nature 402:489-495(1999).  
 CC -1- FUNCTION: INVOLVED IN CYTOKINESIS (POTENTIAL).  
 CC -1- SUBUNIT: MAY ASSEMBLE INTO A MULTICOMPONENT STRUCTURE.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=A;  
 CC IsoId=Q9UH03-1; Sequence=Displayed;  
 CC Name=B;  
 CC IsoId=Q9UH03-2; Sequence=VSP\_006049;  
 CC -1- SIMILARITY: BELONGS TO THE SEPTIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF285107; AAC00517.1; -;  
 CC EMBL: AF285109; AAC00519.1; -;  
 CC EMBL: Z99716; CAB41235.2; ALT\_SEQ.  
 CC PIR: JCT681; JCT681.  
 CC Genew: HGNC:10750; SEPT3.  
 CC GO: GO:0016288; P:cytokinesis; NAS.  
 CC InterPro: IPR000038; GTP\_Cell\_Div.  
 CC Pfam: PF00735; GTP\_CDC; I.  
 CC ProDom: PD002565; GTP\_Cell\_Div; 1.  
 CC Cell division; GTP-binding; Alternative splicing.

FT NP BIND 55 62 GTP (POTENTIAL).  
 FT VARSPIC 325 345 GEGIGTGLPVPVPTCPPTAE -> VSVDESHSNP  
 FT (in isoform B).  
 FT /Ftrid=VSP\_006049.  
 SQ SEQUENCE 345 AA; 39345 MW; B4B6ACAECA60136F CRC64;  
 Query Match 51.4%; Score 36; DB 1; Length 345;  
 Best Local Similarity 55.6%; Pred. No. 42;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ITHIMES 9  
 Db 302 VTHNIHYET 310  
 RESULT 26  
 ID ALF2\_RHOSH STANDARD; PRT; 354 AA.  
 AC P29271;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fructose-bisphosphate aldolase II (EC 4.1.2.13).  
 GN CFXB.  
 OS Rhodospirillum rubrum (Rhodospirillum rubrum).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;  
 CC Rhodospirillaceae; Rhodospirillum.  
 CX NCBI\_TaxID=1063;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9204181; PubMed=1939098;  
 RA Chen J.-H., Gibson J.L., McCue L.A., Tabita F.R.;  
 RT "Identification, expression, and deduced primary structure of  
 RT transketolase and other enzymes encoded within the form II CO2  
 RT fixation operon of Rhodospirillum rubrum";  
 RL J. Biol. Chem. 266:20447-20452(1991).  
 CC -1- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glyceralone  
 CC phosphate + D-glyceraldehyde 3-phosphate.  
 CC -1- COFACTOR: ZINC.  
 CC -1- PATHWAY: GLYCOLYSIS; sixth step.  
 CC -1- PATHWAY: PART OF REDUCTIVE PENTOSE PHOSPHATE PATHWAY OR CALVIN  
 CC CYCLE OF PHOTOSYNTHETIC CARBON DIOXIDE ASSIMILATION.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- MISCELLANEOUS: THIS PROTEIN IS ENCODED WITHIN THE FORM II  
 CC RIBULOSE-BISPHOSPHATE CARBOXYLASE OPERON.  
 CC -1- SIMILARITY: Belongs to class II fructose-bisphosphate aldolase  
 CC family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M68914; AAA26157.1; -;  
 CC PIR: D41080; D41080.  
 CC InterPro: IPR006412; Fruct\_bisp\_Calv.  
 CC InterPro: IPR000771; K\_bp\_aldolase.  
 CC Pfam: PF01116; F\_bp\_aldolase; 1.  
 CC ProDom: PD002376; K\_bp\_aldolase; 1.  
 CC TIGRFAMs: TIGR00167; K\_bp\_aldolase; 1.  
 CC TIGRFAMs: TIGR01521; FruBisAlDO II B. 1.  
 CC PROSITE: PS00602; ALDOASE\_CLASS\_II\_1; 1.  
 CC PROSITE: PS00806; ALDOASE\_CLASS\_II\_2; 1.  
 CC Lysase; Glycolysis; Zinc; Calvin cycle; Multigene family.  
 FT METAL 81 81 ZINC (BY SIMILARITY).  
 FT METAL 84 84 ZINC (BY SIMILARITY).  
 SQ SEQUENCE 354 AA; 38269 MW; 9F547B84C72ACF5 CRC64;  
 Query Match 51.4%; Score 36; DB 1; Length 354;  
 Best Local Similarity 41.7%; Pred. No. 43;

Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 1 ITHIHESASL 12  
Db 128 VSHMAHWGASV 139

RESULT 27  
SEPI\_MOUSE  
ID SEPI\_MOUSE STANDARD; PRT; 465 AA.

AC Q92155;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Neuronal-specific septin 3.  
GN SEPT3 OR SEPI3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Xiong J.-W., Stuhlmann H.;  
RT "Promoter-trap insertion into a novel mammalian septin gene expressed during mouse neuronal development."  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: INVOLVED IN CYTOKINESIS (POTENTIAL).  
CC -1- SUBUNIT: MAY ASSEMBLE INTO A MULTICOMPONENT STRUCTURE.  
CC -1- SIMILARITY: BELONGS TO THE SEPTIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF104411; AAD02884.1; -  
DR MGD; MG11345148; Sept3.  
DR InterPro: IPR000038; GTP\_Cell\_Div.  
DR Pfam: PF00735; GTP\_CDC; 1.  
DR ProDom: PD002565; GTP\_Cell\_Div; 1.  
KM Cell division: GTP-binding  
FT NP BIND 55 GTP (POTENTIAL).  
SQ SEQUENCE 465 AA; 52791 MW; 13B89DC80971382A CRC64;

Query Match 51.4%; Score 36; DB 1; Length 465;  
Best Local Similarity 55.6%; Pred. No. 58;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ITHIHESASL 9  
Db 293 VTHNIHYET 301

RESULT 28  
GGT\_SYNY3  
ID GGT\_SYNY3 STANDARD; PRT; 518 AA.

AC P74181;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Gamma-glutamyltranspeptidase precursor (EC 2.3.2.2).  
GN GGT OR SUR1269.  
OS Synecocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=97061201; PubMed=8905231;  
RA Kaneo T., Sato S., Kocani H., Tanaka A., Asamizu E., Nakamura Y., Miyajima N., Hirosewa M., Sugura M., Sasamoto S., Kimura T.,

RA Hoesouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,  
RA Okumura S., Shimo S., Takeuchi C., Wada T., Watanabe A.,  
RA Yamada M., Yasuda M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions."  
RL DNA Res. 3:109-136(1996).  
CC -1- CATALYTIC ACTIVITY: (5-L-glutamyl)-peptide + an amino acid =  
CC peptide + 5-L-glutamyl-L-amino acid  
CC -1- PATHWAY: GGT PLAYS A KEY ROLE IN THE GAMMA-GLUTAMYL CYCLE, A  
CC PATHWAY FOR THE SYNTHESIS AND DEGRADATION OF GLUTATHIONE.  
CC -1- SUBUNIT: THIS ENZYME CONSISTS OF TWO POLYPEPTIDE CHAINS, WHICH ARE  
CC SYNTHESIZED IN PRECURSOR FORM FROM A SINGLE POLYPEPTIDE.  
CC -1- SIMILARITY: STRONG. TO OTHER GGT ENZYMES AND TO GL-TACA ACYLASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D90913; BA18270.1; -  
DR PIR: S75811; S75811.  
DR MEROPS: T03.013; -  
DR InterPro: IPR000101; Gglutrnaptidase.  
DR Pfam: PF01019; G\_glu\_transpept; 1.  
DR PRINTS: PRO1210; GGTTRANSPASE.  
DR TIGRFAMs: TIGR00066; G\_glu\_trans; 1.  
DR PROSITE: PS00462; G\_GLU\_TRANSPSEPTIDASE; FALSE NEG.  
KM Transferrase; Acyltransferase; zymogen; Signal;  
KM Glutathione biosynthesis; Complete proteome.  
FT SIGNAL 1 336  
FT CHAIN 1 336  
FT CHAIN 337 518  
FT CHAIN 518 518  
SQ SEQUENCE 518 AA; 55988 MW; 4FDA234F16C6540 CRC64;

Query Match 51.4%; Score 36; DB 1; Length 518;  
Best Local Similarity 66.7%; Pred. No. 65;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 4 RIHWESASL 12  
Db 447 RIHWESASL 455

RESULT 29  
U171\_HUMAN  
ID U171\_HUMAN STANDARD; PRT; 569 AA.

AC Q12980; Q92469;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE CgTHBA protein (-14 gene protein).  
GN CgTHBA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=96121379; PubMed=8575760;  
RA Vyas P., Vickers M.A., Picketts D.J., Higgs D.;  
RT "Conservation of position and sequence of a novel, widely expressed  
RT gene containing the major human alpha-globin regulatory element."  
RL Genomics 29:679-689(1995).  
RN [2]  
RP SEQUENCE OF 1-62 FROM N.A.  
RA Thomas K.;  
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.

```

CC -1- TISSUE SPECIFICITY: Widely expressed.
CC -1- SIMILARITY: BELONGS TO THE UPF0171 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC -----
CC EMBL; X90857; CAA62368.1; -
CC EMBL; Z69666; CAA93521.1; -
CC MIM; 600928; -
CC InterPro; IPR005365; UPF0171.
CC Pfam; PF03666; UPF0171.1.
CC SEQUENCE 569 AA; 63604 MW; 44BEF42AA7F2841D CRC64;
SO
Query Match 51.4%; Score 36; DB 1; Length 569;
Best Local Similarity 38.5%; Pred. No. 72;
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 ITHRIHMEASL 13
Db 231 LPHKIHVAASSL 243

RESULT 30
GLMS_BUCAP STANDARD; PRT; 607 AA.
AC Q8KA75;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glucosamine--fructose-6-phosphate aminotransferase [isomerizing]
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (GPA) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
GN GLMS OR BUSG027.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22084549; PubMed=12089438;
RA Tamas I., Klaason L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
RA Wernegreen J.J., Sandstrom J.P., Moran N.A., Andersson S.G.B.;
RT "50 million years of genomic stasis in endosymbiotic bacteria.";
RL Science 296:2376-2379 (2002).
CC -1- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC converting fructose-6P into glucosamine-6P using glutamine as a
CC nitrogen source (By similarity).
CC -1- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC glutamate + D-glucosamine 6-phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SIS FAMILY.
CC GPA7 SUBFAMILY.
CC -1- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC -----
CC EMBL; AE014078; AAM67598.1; -
CC HAMAP; MF_00164; -; 1.
CC InterPro; IPR000583; GATase_2.
CC InterPro; IPR005855; GLMS.
CC InterPro; IPR001347; SIS.
CC Pfam; PF00310; SIS; 2.

```

```

DR Pfam; PF00310; GATase_2; 1.
DR Pfam; PF01380; SIS; 2.
DR TIGRFAMs; TIGR01135; glms; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
KM Transferrase; Aminotransferase; Glutamine amidotransferase;
KW Complete proteome.
FT INT MET 0 BY SIMILARITY.
FT DOMAIN 1 241 GLUTAMINE AMIDOTRANSFERASE.
FT ACT SITE 1 1 GATASE (BY SIMILARITY).
FT ACT_SITE 602 602 ISOMERIZATION FRU-6P (BY SIMILARITY).
SO SEQUENCE 607 AA; 68346 MW; ASC29A5CG03C02B0 CRC64;
SQ
Query Match 51.4%; Score 36; DB 1; Length 607;
Best Local Similarity 62.5%; Pred. No. 78;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 ITHRIHME 8
Db 127 IAHLLHME 134

RESULT 31
GLMS_BUCAI STANDARD; PRT; 608 AA.
AC P57138;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glucosamine--fructose-6-phosphate aminotransferase [isomerizing]
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (GPA) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
GN GLMS OR BU026.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS.";
RL Nature 407:81-86 (2000).
CC -1- FUNCTION: CATALYZES THE FIRST STEP IN HEXOSAMINE METABOLISM.
CC CONVERTING FRUCTOSE-6P INTO GLUCOSAMINE-6P USING GLUTAMINE AS A
CC NITROGEN SOURCE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC glutamate + D-glucosamine 6-phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SIS FAMILY.
CC GPA7 SUBFAMILY.
CC -1- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC -----
CC EMBL; AP001118; BAB12753.1; ALT_INIT.
CC HSSP; P17169; IMQO.
CC MEROPS; C44.001; -
CC HAMAP; MF_00164; -; 1.
CC InterPro; IPR000583; GATase_2.
CC InterPro; IPR005855; GLMS.
CC InterPro; IPR001347; SIS.
CC Pfam; PF00310; GATase_2; 1.
CC Pfam; PF01380; SIS; 2.

```

```

DR  TIGRFAMs; TIGR01135; glms; 1.
DR  PROSITE; PS00443; GATASE_TYPE_II; 1.
KW  Transferase; Aminotransferase; Glutamine amidotransferase;
KW  Complete proteome.
FT  INIT_MET 0
FT  ACT_SITE 1
FT  ACT_SITE 603
FT  DOMAIN 1
FT  SEQUENCE 608 AA; 68550 MW; 77A45F96C8BD190BC CRC64;
SO  SEQUENCE

Query Match
Best Local Similarity 51.4%; Score 36; DB 1; Length 608;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ITHRIHWE 8
Db 127 IATHLHWE 134

RESULT 32
T9S4 HUMAN STANDARD; PRT; 625 AA.
AC Q92544; Q9NUN3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transmembrane 9 superfamily protein member 4.
GN TM9SF4 OR KIAA0255.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=97191544; PubMed=9039502;
RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarayashi Y.,
RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. VI.
RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
RT analysis of cDNA clones from cell line KG-1 and Brain.";
RL DNA Res. 3:321-329(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11790052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam J.R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levenshain M.H., Leverisa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McIlroy K., McKusick A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramay H.,
RA Rice C.M., Ross M.T., Scott C.B., Senta H.K., Shownkeen R., Sims S.,
RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.B.,
RA Swann R.M., Symcote N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [3]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Diatchenko L., Marisano K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Wosley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smallie D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: BELONGS TO THE NONASPANIN (TM9SF) FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
CC -----
DR EMBL; D87444; BAAL3385.1; -
DR EMBL; AL049539; CAB75607.2; -
DR EMBL; BC021107; AAH21107.1; -
DR EMBL; BC022850; AAH22850.1; -
DR InterPro; IPR004240; EMP70.
DR Pfam; PF02990; EMP70; 1.
KW Transmembrane.
FT TRANSMEM 265 285 POTENTIAL.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 360 380 POTENTIAL.
FT TRANSMEM 400 420 POTENTIAL.
FT TRANSMEM 433 453 POTENTIAL.
FT TRANSMEM 485 505 POTENTIAL.
FT TRANSMEM 519 539 POTENTIAL.
FT TRANSMEM 554 574 POTENTIAL.
FT TRANSMEM 582 602 POTENTIAL.
SO SEQUENCE 625 AA; 72541 MW; 69460C8C7E2BEEB5 CRC64;

Query Match
Best Local Similarity 51.4%; Score 36; DB 1; Length 625;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 2 THRIHWEASL 12
Db 234 TYSVHWEESDI 244

RESULT 33.
VAV_RAT STANDARD; PRT; 843 AA.
ID VAV_RAT
AC P54100;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Vav proto-oncogene (p95).
GN VAV1 OR VAV.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```

```

OX NCB1_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99323974; PubMed=10395673;
RA Song J.S., Hafeem-Smith H., Arudchandran R., Gomez J., Scott P.M.,
RA Mill J.F., Tan T.-H., Rivera J.;
RT "Tyrosine phosphorylation of Vav stimulates IL-6 production in mast
RT cells by a Rac/c-Jun N-terminal kinase-dependent pathway.";
RL J. Immunol. 163:802-810(1999).
CC -1- FUNCTION: Couples tyrosine kinase signals with the activation of
CC the Rho/Rac GTPases, thus leading to cell differentiation and/or
CC proliferation.
CC -1- SUBUNIT: Interacts with SLA (By similarity).
CC -1- PTM: Phosphorylated on tyrosine residues.
CC -1- SIMILARITY: Contains 1 calponin-homology (CH) domain.
CC -1- SIMILARITY: Contains 1 DBL-homology (DB) domain.
CC -1- SIMILARITY: Contains 1 PH domain.
CC -1- SIMILARITY: Contains 1 zinc-dependent phorbol-ester and DAG
CC binding domain.
CC -1- SIMILARITY: Contains 1 SH2 domain.
CC -1- SIMILARITY: Contains 2 SH3 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39476; AAA98606.1; -.
DR HSSB; P29354; IGRI.
DR InterPro; IPR001715; Calponin-like.
DR InterPro; IPR003247; CH type.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR001331; GDS_CDCC24.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000219; RhogEF.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR003096; SM22_calponin.
DR Pfam; PF00307; CH; 1.
DR Pfam; PF00130; DAG_PE-bind; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00621; RhogEF; 1.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 2.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00888; SM22CALPONIN.
DR ProDom; PD001527; CH type; 1.
DR ProDom; PD000093; SH2; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00109; CL; 1.
DR SMART; SM00033; CH; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00325; RhogEF; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS50021; CH; 1.
DR PROSITE; PS50479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS50010; DH_2; 1.
DR PROSITE; PS00741; DH_1; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 2.
KW Proto-oncogene; Phorbol-ester binding; Zinc; SH2 domain; SH3 domain;
KW Guanine-nucleotide releasing factor; Repeat; Phosphorylation.
FT DOMAIN 1 119 CH.
FT DOMAIN 194 373 DH.
FT DOMAIN 402 504 PH.
FT DOMAIN 516 564 PHORBOL-ESTER AND DAG BINDING.

```

```

FT DOMAIN 615 658 SH3 1.
FT DOMAIN 669 763 SH2.
FT DOMAIN 780 840 SH3 2.
SQ SEQUENCE 843 AA; 97953 MW; CAA5CAGD45FCB80E CRC64;
Query Match 51.4%; Score 36; DB 1; Length 843;
Best local similarity 45.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Oy 2 THRHMESASL 12
Db 20 SHRYMEGAQV 30
RESULT 34
VAV_MOUSE STANDARD; PRT; 845 AA.
ID_VAV_MOUSE
AC P27870;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Vav proto-oncogene.
GN VAV1 OR VAV.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9228488; PubMed=1565462;
RA Adams J.A., Houston H., Allen J., Lints T., Harvey R.;
RT "The hematopoietically expressed vav proto-oncogene shares homology
RT with the dbl GDP-GTP exchange factor, the bcr gene and a yeast gene
RL (CDC24) involved in cytoskeletal organization.";
RL Oncogene 7:611-618(1992).
RN [2]
RP SEQUENCE OF 1-93 FROM N.A.
RX MEDLINE=91172176; PubMed=2005887;
RA Katrav S., Cleveland J.L., Heeslop H.E., Pulido D.;
RT "Loss of the amino-terminal helix-loop-helix domain of the vav proto-
RL oncogene activates its transforming potential.";
RL Mol. Cell. Biol. 11:1912-1920(1991).
RN [3]
RP INTERACTION WITH SLA.
RX MEDLINE=20130290; PubMed=10662792;
RA Sosnowski T., Pandey A., Dixit V.M., Weiss A.;
RT "Src-like adaptor protein (SLAP) is a negative regulator of T cell
RT receptor signaling.";
RL J. Exp. Med. 191:463-474(2000).
CC -1- FUNCTION: Couples tyrosine kinase signals with the activation of
CC the Rho/Rac GTPases, thus leading to cell differentiation and/or
CC proliferation.
CC -1- SUBUNIT: Interacts with SLA.
CC -1- TISSUE SPECIFICITY: Widely expressed in hematopoietic cells but
CC not in other cell types.
CC -1- PTM: Phosphorylated on tyrosine residues (By similarity).
CC -1- SIMILARITY: Contains 1 calponin-homology (CH) domain.
CC -1- SIMILARITY: Contains 1 DBL-homology (DB) domain.
CC -1- SIMILARITY: Contains 1 PH domain.
CC -1- SIMILARITY: Contains 1 zinc-dependent phorbol-ester and DAG
CC binding domain.
CC -1- SIMILARITY: Contains 1 SH2 domain.
CC -1- SIMILARITY: Contains 2 SH3 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X64361; CAA45713.1; -.

```

DR EMBL; M59833; AAA63402.1; -.  
 DR PIR; A61187; TVMSV.  
 DR PDB; 1F5X; 15-SEP-00.  
 DR PDB; 1GCP; 28-JUN-03.  
 DR PDB; 1GQJ; 28-JAN-03.  
 DR PDB; 1K1Z; 18-DEC-02.  
 DR TRANSFAC; T01230; -.  
 DR MGD; MGI:98923; Vav1.  
 DR InterPro; IPR001715; Calponin-like.  
 DR InterPro; IPR003247; CH\_type.  
 DR InterPro; IPR002219; DAG\_pe-bind.  
 DR InterPro; IPR001331; GDS\_CDC24.  
 DR InterPro; IPR001849; PH\_.  
 DR InterPro; IPR000219; RhogEF.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR003096; SM22\_calponin.  
 DR Pfam; PF00307; CH; 1.  
 DR Pfam; PF00130; DAG\_pe-bind; 1.  
 DR Pfam; PF00169; PH; 1.  
 DR Pfam; PF00621; RhogEF; 1.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 2.  
 DR PRINTS; PR00401; SH2DOMAIN.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR PRINTS; PR00888; SM22CALPONIN.  
 DR ProDom; PD001527; CH type; 1.  
 DR ProDom; PD000093; SH2; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00109; C1; 1.  
 DR SMART; SM00033; CH; 1.  
 DR SMART; SM00033; PH; 1.  
 DR SMART; SM00325; RhogEF; 1.  
 DR SMART; SM00325; SH2; 1.  
 DR SMART; SM00326; SH3; 2.  
 DR PROSITE; PS50021; CH; 1.  
 DR PROSITE; PS00479; DAG\_pe BIND DOM 1; 1.  
 DR PROSITE; PS50081; DAG\_pe BIND\_DOM\_2; 1.  
 DR PROSITE; PS50010; DH\_2; 1.  
 DR PROSITE; PS00741; DH\_1; 1.  
 DR PROSITE; PS50003; PH\_DOMAIN; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 DR PROSITE; PS50002; SH3; 2.  
 KM Proto-oncogene; Phorbol-ester binding; Zinc; SH2 domain; SH3 domain;  
 KW 3D-structure.  
 KM Guanine-nucleotide releasing factor; Repeat; Phosphorylation;  
 FT DOMAIN 1 119 CH.  
 FT DOMAIN 194 373 DH.  
 FT DOMAIN 402 504 PH.  
 FT DOMAIN 516 564 PHORBOL-ESTER AND DAG BINDING.  
 FT DOMAIN 617 660 SH3 1.  
 FT DOMAIN 671 765 SH2.  
 FT DOMAIN 782 842 SH3 2.  
 FT CONFLICT 29 Q -> E (IN REF. 2).  
 SQ SEQUENCE 845 AA; 98136 MW; 3666DCD1C5229DA CRC64;

Query Match 51.4%; Score 36; DB 1; Length 845;  
 Best Local Similarity 45.5%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

DR Db 2 THRIHMSASL 12  
 20 SHRYTWGAGV 30

RESULT 35  
 ID VAV3 HUMAN STANDARD; PRT; 847 AA.  
 AC O9UKM4; O95230; O9Y5X8.  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Vav-3 protein.

GN VAV3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.  
 RX MEDLINE=9455043; PubMed=10523675;  
 RA Mavilla N., Bustelo X.R.,  
 RT "Biological and regulatory properties of Vav-3, a new member of the  
 RT Vav family of oncoproteins."  
 RL Mol. Cell. Biol. 19:7870-7885(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast, and Colon carcinoma;  
 RX MEDLINE=98371222; PubMed=9705494;  
 RA Treinkle T., Welsh J., Jung B., Mathieu-Daude F., McClelland M.,  
 RT "Non-stoichiometric reduced complexity probes for cDNA arrays."  
 RL Nucleic Acids Res. 26:3883-3891(1998).  
 CC -1- FUNCTION. EXCHANGE FACTOR FOR GTP-BINDING PROTEINS RHOA, RHOG AND  
 TO A LESSER EXTENT, RAC-1. BINDS PHYSICALLY TO THE NUCLEOTIDE-FREE  
 STATES OF THOSE GTPASES.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Alpha;  
 CC IsoId=O9UKM4-1; Sequence=Displayed;  
 CC Name=Beta;  
 CC IsoId=O9UKM4-2; Sequence=VSP\_001820;  
 CC -1- SIMILARITY: Contains 1 calponin-homology (CH) domain.  
 CC -1- SIMILARITY: Contains 1 DBL-homology (DH) domain.  
 CC -1- SIMILARITY: Contains 1 PH domain.  
 CC -1- SIMILARITY: Contains 1 zinc-dependent phorbol-ester and DAG  
 binding domain.  
 CC -1- SIMILARITY: Contains 1 SH2 domain.  
 CC -1- SIMILARITY: Contains 2 SH3 domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).  
 CC -----

DR EMBL; AF118887; AAP20349.1; -.  
 DR EMBL; AF118886; AAD20348.1; -.  
 DR EMBL; AF067817; AAC79695.1; -.  
 DR HSSP; P29355; ISEM.  
 DR Genew; HGNC:12659; VAV3.  
 DR MIM; 605541; -.  
 DR GO; GO:0005096; F:GTPase activator activity; TAS.  
 DR GO; GO:0005070; F:SH3/SH2 adaptor protein activity; TAS.  
 DR GO; GO:0007264; P:small GTPase mediated signal transduction; TAS.  
 DR InterPro; IPR001715; Calponin-like.  
 DR InterPro; IPR003247; CH type.  
 DR InterPro; IPR002219; DAG\_pe-bind.  
 DR InterPro; IPR001331; GDS\_CDC24.  
 DR InterPro; IPR001849; PH.  
 DR InterPro; IPR000219; RhogEF.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR003096; SM22\_calponin.  
 DR Pfam; PF00307; CH; 1.  
 DR Pfam; PF00130; DAG\_pe-bind; 1.  
 DR Pfam; PF00169; PH; 1.  
 DR Pfam; PF00621; RhogEF; 1.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR00401; SH2DOMAIN.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR PRINTS; PR00888; SM22CALPONIN.  
 DR ProDom; PD001527; CH type; 1.  
 DR ProDom; PD000093; SH2; 1.





RESULT 37  
 CO3\_RAT STANDARD: PRT: 1663 AA.  
 AC P01026;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Complement C3 precursor [Contains: C3a anaphylatoxin].  
 GN C3.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Miscari; TISSUE=Liver;  
 RX MEDLINE=90245672; PubMed=2336397;  
 RA Mitsuami Y., Sohma M., Ikehara Y.;  
 RT "Nucleotide and deduced amino acid sequence of rat complement C3";  
 RL Nucleic Acids Res. 18:2178-2178(1990).  
 RN [2]  
 RP SEQUENCE OF 671-748.  
 RX MEDLINE=79062262; PubMed=309768;  
 RA Jacobs J.W., Rubin J.S., Hugli T.E., Bogardt R.A., Mariz I.K.,  
 RT "Purification, characterization, and amino acid sequence of rat anaphylatoxin (C3a)";  
 RL Biochemistry 17:5031-5038(1978).  
 RN [3]  
 RP SEQUENCE OF 1316-1595 FROM N.A.  
 RX MEDLINE=89380332; PubMed=2674144;  
 RA Sundstrom S.A., Komm B.S., Ponce-De-Leon H., Yi Z., Teuscher C.,  
 RT "Regulation of tissue-specific expression of complement C3";  
 RL J. Biol. Chem. 264:16941-16947(1989).  
 CC -1- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS. AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE THIOESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES. C3A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND BASOPHILIC LEUKOCYTES.  
 CC -1- SUBUNIT: C3 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 ARG RESIDUES, FORMING TWO CHAINS, BETA & ALPHA, LINKED BY A DISULFIDE BOND. C3 CONVERTASE ACTIVATES C3 BY CLEAVING THE ALPHA CHAIN, RELEASING C3A ANAPHYLATOXIN & GENERATING C3B (BETA CHAIN + ALPHA CHAIN).  
 CC -1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.  
 CC -1- SIMILARITY: Contains 1 anaphylatoxin-like domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL; X52477; CAA36716.1; -  
 CC EMBL; M29866; AAA40837.1; ALT\_SEQ.  
 CC PIR; S15764; C3RT.  
 CC PDB; 1QOF; 31-JUL-00.  
 CC InterPro; IPR002890; A2M\_N.  
 CC InterPro; IPR000020; Anaphylatoxin.  
 CC InterPro; IPR001599; Macroglobulin2.  
 CC InterPro; IPR01134; Netrin\_C.  
 CC Pfam; PF00207; A2M; 1.  
 CC Pfam; PF01835; A2M\_N; 1.

DR Pfam; PF01821; ANATO; 1.  
 DR Pfam; PF01759; NTR; 1.  
 DR ProDom; PD003264; Anaphylatoxin; 1.  
 DR SMART; SM00104; ANATO; 1.  
 DR SMART; SM00643; C345C; 1.  
 DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 DR PROSITE; PS01177; ANAPHYLATOXIN\_1; 1.  
 DR PROSITE; PS01178; ANAPHYLATOXIN\_2; 1.  
 KW Complement pathway; Complement alternate pathway; Plasma; Inflammatory response; Glycoprotein; Signal; 3D-structure;  
 KW Thioester bond.  
 FT SIGNAL 1 24  
 FT CHAIN 25 1663 COMPLEMENT C3.  
 FT CHAIN 25 666 BETA CHAIN.  
 FT CHAIN 671 1663 ALPHA CHAIN.  
 FT PEPTIDE 671 748 C3A ANAPHYLATOXIN.  
 FT CHAIN 749 1663 C3B (ALPHA CHAIN).  
 FT SITE 748 749 CLEAVAGE (BY C3 CONVERTASE).  
 FT DOMAIN 693 728 ANAPHYLATOXIN-LIKE.  
 FT DISULFID 558 816 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 626 661 BY SIMILARITY.  
 FT DISULFID 693 720 BY SIMILARITY.  
 FT DISULFID 694 727 BY SIMILARITY.  
 FT DISULFID 707 728 BY SIMILARITY.  
 FT DISULFID 873 1513 BY SIMILARITY.  
 FT DISULFID 1101 1158 BY SIMILARITY.  
 FT DISULFID 1358 1489 BY SIMILARITY.  
 FT DISULFID 1389 1458 BY SIMILARITY.  
 FT DISULFID 1506 1511 BY SIMILARITY.  
 FT DISULFID 1518 1590 BY SIMILARITY.  
 FT DISULFID 1537 1661 BY SIMILARITY.  
 FT CROSSLINK 1010 1013 Iso-glutamy1 cysteine thioester (Cys-Gln).  
 FT CARBOHYD 939 939 N-LINKED (GLCNAC. . .) (PROBABLE).  
 FT CONFLICT 721 722 LK -> KL (IN REF. 2).  
 SQ SEQUENCE 1663 AA; 186460 MW; 2F87CCB143CD4BC CRC64;  
 Query Match 51.4%; Score 36; DB 1; Length 1663;  
 Query Local Similarity 70.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 RHIMESASLT 13  
 DB 1310 RLMESSGL 1319  
 RESULT 38  
 POLG\_TBRES STANDARD: PRT: 3412 AA.  
 AC P07720; P07721; Q88475; Q88476; Q88477; Q88478; Q88479; Q88877;  
 AC Q88878; Q88879;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Genome polypeptide [Contains: Capsid protein C (core protein); Matrix protein; Envelope protein (M); Major envelope protein E; Nonstructural proteins NS1, NS2A, NS2B, NS4A and NS4B; Protease/helicase (EC 3.4.21.98) (NS3); RNA-directed RNA polymerase (EC 2.7.7.48) (NS5)].  
 DE (EC 3.4.21.98) (NS3); RNA-directed RNA polymerase (EC 2.7.7.48) (NS5)].  
 OS Tick-borne encephalitis virus (strain Sofjin) (TBEV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Flavivirus.  
 OX NCBI\_TaxID=11087;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90101381; PubMed=2136778;  
 RA Plotnev A.G., Vamshchikov V.F., Blinov V.M.;  
 RT "Nucleotide sequence of tick-borne encephalitis virus. I; complete amino acid sequence of the polypeptide of tick-borne encephalitis virus.";  
 RL Virology 174:250-263(1990).  
 RN [2]  
 RP SEQUENCE OF 1-1190 FROM N.A.  
 RX MEDLINE=88319988; PubMed=2970626;

RA Yamshchikov V.F., Pletnev A.G.;  
 RT "Nucleotide sequence of the genome region encoding the structural  
 RT proteins and the NS1 protein of the tick borne encephalitis virus.";  
 RL Nucleic Acids Res. 16:7750-7750(1988).  
 RN [3]  
 RP SEQUENCE OF 1-683 AND 777-1002 FROM N.A.  
 RX MEDLINE=86220766; PubMed=3709796;  
 RA Pletnev A.G., Yamshchikov V.F., Blinov V.M.;  
 RT "Tick-borne encephalitis virus genome. The nucleotide sequence coding  
 RT for virion structural proteins.";  
 RL FEBS Lett. 200:317-321(1986).  
 CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE  
 CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.  
 CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral  
 CC precursor polypeptide, commonly with Asp or Glu in the p6  
 CC position, Cys or Thr in p1 and Ser or Ala in p1'.  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC {RNA} (N).  
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A  
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:  
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  
 CC PROTEIN C AND RNA.  
 CC -1- MISCELLANEOUS: THE NONSTRUCTURAL PROTEINS NS1 PRESENTS TWO  
 CC ALTERNATIVE CLEAVAGE SITES FOR ITS C-TERMINUS, WHICH MAY DEFINE A  
 CC SOLUBLE OR A MEMBRANE-BOUND FORM OF NS1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC  
 DR EMBL; X07755; CAA30581.1; -.  
 DR EMBL; X03870; CAA27500.1; -.  
 DR EMBL; X03870; CAA27501.1; ALT\_SEQ.  
 DR EMBL; X03870; CAA27502.1; ALT\_SEQ.  
 DR EMBL; X03870; CAA27503.1; ALT\_SEQ.  
 DR EMBL; X03871; CAA27505.1; -.  
 DR MEROPS; S07.0PW; -.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR001122; Flavi\_capsidC.  
 DR InterPro; IPR000336; Flavi\_glycoprote.  
 DR InterPro; IPR001850; Flavi\_helicase.  
 DR InterPro; IPR000069; Flavi\_M.  
 DR InterPro; IPR001157; Flavi\_NS1.  
 DR InterPro; IPR000752; Flavi\_NS2A.  
 DR InterPro; IPR000487; Flavi\_NS2B.  
 DR InterPro; IPR000404; Flavi\_NS4A.  
 DR InterPro; IPR001528; Flavi\_NS4B.  
 DR InterPro; IPR000208; Flavi\_NS5.  
 DR InterPro; IPR002535; Flavi\_propep.  
 DR InterPro; IPR002877; Flavi.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR InterPro; IPR007095; RNA\_pol\_PS.  
 DR InterPro; IPR007094; RNA\_pol\_psvir.  
 DR Pfam; PF01003; Flavi\_capsid.1.  
 DR Pfam; PF02832; Flavi\_glycop.C; 1.  
 DR Pfam; PF00869; Flavi\_glycoprot; 1.  
 DR Pfam; PF00949; Flavi\_helicase; 1.  
 DR Pfam; PF01004; Flavi\_M; 1.  
 DR Pfam; PF00948; Flavi\_NS1; 1.  
 DR Pfam; PF01005; Flavi\_NS2A; 1.  
 DR Pfam; PF01002; Flavi\_NS2B; 1.  
 DR Pfam; PF01350; Flavi\_NS4A; 1.  
 DR Pfam; PF01349; Flavi\_NS4B; 1.  
 DR Pfam; PF00972; Flavi\_NS5; 1.  
 DR Pfam; PF01570; Flavi\_propep; 1.  
 DR Pfam; PF01728; Flavi\_M; 1.  
 DR Pfam; PF00771; Flavi\_glycoprote; 1.  
 DR Pfam; PF00771; Flavi\_glycoprote; 1.  
 DR Pfam; PF00771; Flavi\_glycoprote; 1.

DR ProDom; PD001496; Flavi\_NS1; 1.  
 KW Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;  
 KW Core protein; Coat protein; Envelope protein; Hydrolyase; Helicase;  
 KW ATP-binding; Transmembrane; Nonstructural protein.  
 FT INT\_MET 1  
 FT CHAIN 1  
 FT PROPEP 113 112  
 FT CHAIN 206 280  
 FT CHAIN 281 776  
 FT CHAIN 777 71128  
 FT CHAIN 71129 1358  
 FT CHAIN 1359 1489  
 FT CHAIN 1490 2110  
 FT CHAIN 2111 2259  
 FT CHAIN 2260 2510  
 FT CHAIN 2511 3412  
 FT NP\_BIND 1688 1695  
 FT SITE 1779 1782  
 FT TRANSMEM 101 112  
 FT TRANSMEM 247 259  
 FT TRANSMEM 266 280  
 FT TRANSMEM 738 751  
 FT DISULFD 283 310  
 FT DISULFD 340 396  
 FT DISULFD 354 385  
 FT DISULFD 372 401  
 FT DISULFD 466 570  
 FT DISULFD 587 618  
 FT CARBOHYD 144 144  
 FT CARBOHYD 434 434  
 FT CARBOHYD 861 861  
 FT CARBOHYD 983 983  
 FT CARBOHYD 999 999  
 FT CARBOHYD 1228 1228  
 FT CARBOHYD 2447 2447  
 FT CARBOHYD 2466 2466  
 FT CONFLICT 381 381  
 FT CONFLICT 850 850  
 FT SEQUENCE 3412 AA; 377976 MW; 0f61c6edcdcd965 CRC64;  
 Query Match 51.4%; Score 36; DB 1; Length 3412;  
 Best Local Similarity 71.4%; Pred. No. 5.1e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 HRIMES 9  
 Db 3399 HDLMHES 3405  
 RESULT 39  
 WTI1\_HUMAN STANDARD; PRT; 92 AA.  
 AC Q06250; Q96A27;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Wlms' tumor associated protein (WT1-1).  
 GN WTI1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RA MEDLINE=94010952; PubMed=8406502;  
 RA Gaessler M., Bruns G.A.P.;  
 RT "Sequence of the WTI upstream region including the Wti-1 gene.";  
 RL Genomics 17:499-501(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;

```

RX MEDLINE=91048012; PubMed=2173145;
RA Huang A., Campbell C.E., Bonetta L., McAndrews-Hill M.S.,
RA Chilton-MacNeill S., Coppes M.J., Law D.J., Feinberg A.P.,
RA Yeager H., Williams B.R.;
RT "tissue, developmental, and tumor-specific expression of divergent
RT transcripts in Wilms tumor.";
RL Science 250:991-994(1990).
RN [3]
RP SEQUENCE FROM N.A.
RA Frankland J.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshnyki S., Carrinci P., Frange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bock S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Watling M., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69950; CAA49572.1; -
DR EMBL; M60614; AAA63269.1; -
DR EMBL; AL049692; CAC39219.1; -
DR EMBL; BC002734; AA02734.1; -
DR PIR; A36233; A36233.
DR PIR; I39287; I39287.
DR GO; GO:0006283; P:cell proliferation; NMS.
KW Polymorphism.
FT VARIANT 13 13 A -> T (in dbSNP:6508).
FT /FTID=VAR_014631.
FT CONFLICT 83 92 SGNHPPAADN -> PGVQRPRTTE (IN REF. 1).
FT SEQUENCE 92 AA; 10038 MW; 2FE0529B32FAC64 CRC64;
SQ
Query Match 50.0%; Score 35; DB 1; Length 92;
Best Local Similarity 62.5%; Pred. No. 15;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 5 IHMESASL 12
DB 15 IHMQSAGI 22

```

```

RESULT 40
YM71 MYCTU STANDARD; PRT; 99 AA.
AC Q50692.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE Hypothetical protein Ry2271;
GN RV2271 OR MT2332 OR MTCY339.39C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=96295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers R.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickley E.,
RA Kolony J.F., Nelson M.C., Umayan L.A., Ernolsava M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mkhia A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z77163; CAB00964.1; -
DR EMBL; AE007076; AAK46415.1; -
DR PIR; C70730; C70730.
DR TIGR; MT2332; -
DR TubercuList; RV2271; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 99 AA; 11535 MW; 276DA391CE96B85B CRC64;

```

```

Query Match 50.0%; Score 35; DB 1; Length 99;
Best Local Similarity 62.5%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 2 THRIMES 9
DB 65 SHRIEMET 72

```

Search completed: August 28, 2003, 14:10:47  
Job time : 26 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2003, 14:10:13 ; Search time 93 Seconds  
(without alignments)  
36.072 Million cell updates/sec

Title: US-09-845-738c-1  
Perfect score: 70  
Sequence: 1 ITHRIMESASL 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	74.3	154	6 Q29289	Q29289 sus scrofa
2	52	74.3	167	6 Q9NOM4	Q9NOM4 cervus nipp
3	52	74.3	349	6 Q46544	Q46544 ovis aries
4	52	74.3	1661	6 Q9GKPI	Q9GKPI sus scrofa
5	46	65.7	267	16 Q9HRT25	Q9HRT25 pseudomonas
6	42	60.0	211	16 Q9HYZ4	Q9HYZ4 pseudomonas
7	42	60.0	401	16 Q9P9Y5	Q9P9Y5 xylella fas
8	42	60.0	407	5 Q8S1Y7	Q8S1Y7 dirosophila
9	41	58.6	219	13 Q9OYC5	Q9OYC5 brachydanio
10	41	58.6	441	5 Q8T3J9	Q8T3J9 dirosophila
11	41	58.6	615	16 Q9CHM3	Q9CHM3 lactococcus
12	41	58.6	863	16 Q8DQO6	Q8DQO6 synchococc
13	40	57.1	228	16 Q8Y1Y6	Q8Y1Y6 bruceella su
14	40	57.1	229	16 Q8FTX0	Q8FTX0 mus musculu
15	40	57.1	274	11 Q9D912	Q9D912 mus musculu
16	40	57.1	285	5 Q18611	Q18611 caenorhabdi

17	40	57.1	338	16 Q8PF47	Q8PF47 xanthomonas
18	40	57.1	347	5 Q9VR24	Q9VR24 dirosophila
19	40	57.1	541	16 Q9A017	Q9A017 streptococc
20	40	57.1	541	16 Q8P0P2	Q8P0P2 streptococc
21	40	57.1	1145	11 Q9DBV3	Q9DBV3 mus musculu
22	39	55.7	172	16 Q8PR28	Q8PR28 xanthomonas
23	39	55.7	197	17 Q9HK18	Q9HK18 thermoplasma
24	39	55.7	232	16 Q92KX1	Q92KX1 rhizobium m
25	39	55.7	321	11 Q91ZC0	Q91ZC0 mus musculu
26	39	55.7	322	11 Q91ZB9	Q91ZB9 mus musculu
27	39	55.7	381	2 Q8RTQ7	Q8RTQ7 thermodesul
28	39	55.7	382	2 Q93EV7	Q93EV7 thermodesul
29	39	55.7	387	16 Q8CUT5	Q8CUT5 streptomyce
30	39	55.7	406	2 P95619	P95619 rhodocyclu
31	39	55.7	406	2 Q9JPB4	Q9JPB4 rhodocyclu
32	39	55.7	410	16 Q8FTU8	Q8FTU8 corynebacte
33	39	55.7	437	5 Q8MMP7	Q8MMP7 dictyostell
34	39	55.7	572	16 Q8B3H4	Q8B3H4 streptococc
35	39	55.7	572	16 Q8DXV6	Q8DXV6 streptococc
36	39	55.7	574	16 Q927C5	Q927C5 listeria in
37	39	55.7	574	16 Q8Y3W3	Q8Y3W3 listeria in
38	38	54.3	250	4 Q9NS78	Q9NS78 homo sapien
39	38	54.3	265	4 Q9NS79	Q9NS79 homo sapien
40	38	54.3	266	16 Q8UBV4	Q8UBV4 agrobacteri
41	38	54.3	269	4 Q9H4F5	Q9H4F5 homo sapien
42	38	54.3	280	5 Q45477	Q45477 caenorhabdi
43	38	54.3	317	13 Q9OMC0	Q9OMC0 agkistrodon
44	38	54.3	432	16 Q8G4P1	Q8G4P1 bifidobacte
45	38	54.3	439	5 Q9NEC1	Q9NEC1 leishmania

#### ALIGNMENTS

RESULT 1  
ID Q29289 PRELIMINARY; PRT; 154 AA.  
AC Q29289;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DT Complement C3 (Fragment).  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Small intestine;  
RX MEDLINE=96327607; PubMed=8672129;  
RA Winteroe A.K., Fredholm M., Davies W.,  
RT "Evaluation and characterization of a porcine small intestine cDNA library."  
RL Mamm. Genome 7:509-517(1996).  
DR EMBL; F14640; CAA23173.1; -.  
DR HSSP; P01024; 1C3D.  
DR InterPro; IPR001599; Macrogloblina2.  
DR Pfam; PF00207; AZM; 1.  
FT NON\_TER 1  
FT NON\_TER 154  
SQ SEQUENCE 154 AA; 6DC7661C1253ED45 CRC64;  
Query Match 74.3%; Score 52; DB 6; Length 154;  
Best Local Similarity 76.9%; Pred. No. 0.25;  
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Cy 1 ITHRIMESASL 13  
Db 100 VRRHIMESASL 112  
RESULT 2  
Q9NOM4

```

ID 09N0M4 PRELIMINARY; PRT; 167 AA.
AC 09N0M4;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Complement C3 alpha chain (Fragment).
OS Cervus nippon (Sika deer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
OC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID=9863;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Jiang Y., Sun L.G., Yu Y.L.;
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF264631; AAF73464.1; -.
DR HSSP; P01024; IC3D.
DR InterPro; IPR001599; Macrogloblina2.
DR Pfam; PF00207; A2M; 1.
FT NON TER 1
SQ SEQUENCE 167 AA; 18671 MW; 12BFE0798290DFA7 CRC64;

Query Match 74.3%; Score 52; DB 6; Length 167;
Best Local Similarity 76.9%; Pred. No. 0.27;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

RESULT 3
O46544 PRELIMINARY; PRT; 349 AA.
AC O46544;
DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Complement component C3 (Fragment).
GN C3.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=white alpine; TISSUE=Liver;
RX MEDLINE=98309471; PubMed=9647256;
RA Hein W.R., Dudler L., Marston W.L., Landsverk T., Young A.J.,
RA Avila D.;
RT "Ubiquitination and dimerization of complement receptor type 2 on
RT sheep B cells.";
RL J. Immunol. 161:458-466(1998).
DR EMBL; AF038130; AAB92374.2; -.
DR HSSP; P01024; IC3D.
DR InterPro; IPR001599; Macrogloblina2.
DR Pfam; PF00207; A2M; 1.
DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
FT NON TER 1
FT NON TER 349
SQ SEQUENCE 349 AA; 39679 MW; 70C2023E42ED5EE3 CRC64;

```

```

Query Match 74.3%; Score 52; DB 6; Length 349;
Best Local Similarity 76.9%; Pred. No. 0.59;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 ITHRIWESASL 13
Db 331 VKHRIWESASL 343

```

```

RESULT 4
ID 09GKP1 PRELIMINARY; PRT; 1661 AA.
AC 09GKP1;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Complement component C3.
GN C3.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=21313047; PubMed=11419349;
RA Wimmers K., Mekchay S., Pensuksill S., Hardge T., Yerde M.,
RA Schellander K.;
RT "Polymorphic sites in exon 15 and 30 of the porcine C3 gene.";
RL Anim. Genet. 32:46-47(2001).
DR EMBL; AF154933; AAG40565.1; -.
DR HSSP; P01024; IC3D.
DR InterPro; IPR002890; A2M N.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR001840; Anaphylatoxin.
DR InterPro; IPR001599; Macrogloblina2.
DR InterPro; IPR001134; Netrin_C.
DR Pfam; PF00207; A2M; 1.
DR Pfam; PF01835; A2M N; 1.
DR Pfam; PF01821; ANATO; 1.
DR Pfam; PF01759; NTR; 1.
DR PRINTS; PR000004; ANAPHYLATOXN.
DR ProDom; PD003264; Anaphylatoxin; 1.
DR SMART; SM00104; ANATO; 1.
DR SMART; SM00643; C345C; 1.
DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
SQ SEQUENCE 1661 AA; 186806 MW; 4899D0914BE3310C CRC64;

Query Match 74.3%; Score 52; DB 6; Length 1661;
Best Local Similarity 76.9%; Pred. No. 3;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 ITHRIWESASL 13
Db 1305 VRRHRIWESASL 1317

RESULT 5
ID 09HTZ5 PRELIMINARY; PRT; 267 AA.
AC 09HTZ5;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical protein PA5194.
GN PA5194.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stever C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;

```

RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 DR EMBL: AB004932; AAC08579.1; -  
 DR InterPro: IPR000326; PA\_ptase.  
 DR Pfam: PF01569; PAP2.1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 267 AA; 30527 MW; 57CD9D2319B6AD7D CRC64;

Query Match 65.7%; Score 46; DB 16; Length 267;  
 Best Local Similarity 58.3%; Pred. No. 4.9;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ITRIHESASL 12  
 DB 120 IAHHLHWQHASL 131

RESULT 6  
 Q9HYZ4 PRELIMINARY; PRT; 211 AA.  
 AC Q9HYZ4;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Pseudouridine synthase RluA.  
 GN RLUA OR PA3246.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stever C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wedman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Labbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.,  
 RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.";  
 RT Nature 406:959-964(2000).  
 RL EMBL: AB004747; AAC06634.1; -  
 DR InterPro: IPR006145; Pseudou\_synth.  
 DR InterPro: IPR006224; PSI\_RLU.  
 DR Pfam: PF00849; Pseudou\_synth\_2; 1.  
 DR ProDom: PD001819; PSI\_RLU; 1.  
 DR PROSITE: PS01129; PSI\_RLU; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 211 AA; 24338 MW; D33B20FCEA5A594 CRC64;

Query Match 60.0%; Score 42; DB 16; Length 211;  
 Best Local Similarity 46.2%; Pred. No. 19;  
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 ITRIHESASL 13  
 DB 52 IYHRLDWTSGLM 64

RESULT 7  
 Q9PY5 PRELIMINARY; PRT; 401 AA.  
 AC Q9PY5;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Hypothetical protein Xc2735.  
 GN Xc2735.  
 OS Xylella fastidiosa.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;

OC Xanthomonadaceae; Xylella.  
 OX NCBI\_TaxID=2371;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=9a5c;  
 RX MEDLINE=20365717; PubMed=10910347;  
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
 RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,  
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,  
 RA Colauto N.B., Colombo C., Costa F.P., Costa M.C.R., Costa-Neto C.M.,  
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,  
 RA Facinanci A.P., Ferreira A.J.S., Ferreira V.C.J., Ferro J.A.,  
 RA Fraga J.S., Franca S.C., Franco M.C., Frome M., Furlan L.R.,  
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 RA Ho P.L., Honeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pequeiro J.B.,  
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A., Jr.,  
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Teal S.M., Tsuchioka A.L.,  
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Zeldanis J., Zerbini J.C.,  
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";  
 RL Nature 406:151-159(2000).  
 DR EMBL: AB004080; AAF85520.1; -  
 KM Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 401 AA; 45544 MW; 050ADA9125A6398 CRC64;

Query Match 60.0%; Score 42; DB 16; Length 401;  
 Best Local Similarity 45.8%; Pred. No. 37;  
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ITRIHESASL 11  
 DB 334 LAHRVHDEES 344

RESULT 8  
 Q8SY7 PRELIMINARY; PRT; 407 AA.  
 AC Q8SY7; Q9VA14;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE RE27547p (CG1859 protein).  
 GN CG1859.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champ M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nunco J., Pacleby J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Ceiniker S.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN=Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.E., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer J.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,  
 RA Abri'l J.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.T., Benos P.V., Bereman B.P., Bhandari D., Bolintinas S.,  
 RA Borkova D., Botchan M.A., Bouck J., Brokstein P., Brotler P.,  
 RA Butts K.C., Busam D.A., Butler H., Cadieu L., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies S.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalaal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Pollard J., Puri V., Reese M.G.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA She B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier B., Spradling A.C., Stapleton M., Stroup K., Sun E.,  
 RA Svirskeas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Fairtan D.,  
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jalaal M., Kruse D., Li P., Matel B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Pacleb J., Paragab V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskeas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 [4]  
 RN SEQUENCE FROM N.A.  
 RP Miara S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochick S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Bernan B., Carlson J.W., Celniker S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Kronmiller B., Marshall B., Milburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Slatk J., Shu S., Smutnick F., Whitfield A.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RT "Annotation of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 [5]  
 RN SEQUENCE FROM N.A.  
 RP Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 [6]

RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 DR EMBL; AY071238; AAL48860.1;  
 DR EMBL; AE003842; AAF59286.2;  
 DR FlyBase; FBGN003147; CG1859.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 KM Protease inhibitor; Serine protease inhibitor; Serpin.  
 SQ SEQUENCE 407 AA; 44863 MW; 5D2A46A75C86DD78 CRC64;  
 Query Match 60.0%; Score 42; DB 5; Length 407;  
 Best Local Similarity 77.8%; Pred. No. 37;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 3 HRHMSAS 11  
 Db 150 HRLMSAS 158  
 RESULT 9  
 ID Q90YC5 PRELIMINARY; PRT; 219 AA.  
 AC Q90YC5;  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Ephin-A3.  
 GN EFNA3 OR EPHRIN-A3.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 CC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21412237; PubMed=11520665;  
 RA Hirata Y., Mieda M., Harada T., Yamasu K., Okamoto H.;  
 RT "Identification of ephin-A3 and novel genes specific to the midbrain-  
 MHB in embryonic zebrafish by ordered differential display.";  
 RL Mech. Dev. 107:83-96(2001).  
 DR EMBL; AB051678; BAB55891.1;  
 DR ZFIN; ZDB-GENE-01108-1; efna3.  
 DR InterPro; IPR001799; Ephin.  
 DR Pfam; PF00812; Ephin; 1.  
 DR PRINTS; PR01347; EPHRIN.  
 DR PRODOM; PD002533; Ephin; 1.  
 DR PROSITE; PS01299; EPHRIN; 1.  
 SQ SEQUENCE 219 AA; 25146 MW; 7191927E03F8EA01 CRC64;  
 Query Match 58.6%; Score 41; DB 13; Length 219;  
 Best Local Similarity 45.5%; Pred. No. 29;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 Qy 3 HRHMSAS 13  
 Db 26 HAVHMSAS 36  
 RESULT 10  
 ID Q8T3J9 PRELIMINARY; PRT; 441 AA.  
 AC Q8T3J9; Q9VIX7;  
 DT 01-JUN-2002 (TREMblrel. 21, Created)  
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
 DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)  
 DE ATR1889p (CG7196 protein).  
 GN CG7196.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;



OC Ephydroidea, Drosophilidae; Drosophila.  
 RX NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Lao G.,  
 RA Miranda A., Munzall C.J., Nunco J., Paclel J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celniker S.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Berens P.V., Bernier B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Butts K.C., Buzam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Engelstler C.C., Ferraz C., Ferriere S., Fleischmann W.,  
 RA Fodor C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glöckner A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Honick J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbegam C.,  
 RA Jalaal M., Kalush F., Karpis G.H., Ke Z., Kemison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Krat C., Kravitz S., Kulp D., Lai Z.,  
 RA Laake P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paclel J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banon J., An H., Baldwin D., Banon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Honick J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibbegam C., Jalaal M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,  
 RA Paclel J., Paragas V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of Drosophila melanogaster genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Miera S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,

RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Munzall C.J., Lewis S.E.,  
 RA "Annotation of Drosophila melanogaster genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA FLYBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY094997; AM11325.1; -;  
 DR EMBL: AB003618; AAF52552.2; -;  
 DR FLYBase: FBgn0031944; CG7196.  
 SQ SEQUENCE 441 AA; 52125 MW; 847067D8FA3A3A16 CRC64;  
 QY Query Match 58.6%; Score 41; DB 5; Length 441;  
 DB Best Local Similarity 46.2%; Pred. No. 60;  
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 1 ITRIHVESASLL 13  
 21 VVKHMKRQVSL 33  
 RESULT 11  
 ID 09GCHM3 PRELIMINARY; PRT; 615 AA.  
 AC 09GCHM3;  
 DT 01-JUN-2001 (TEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)  
 DE Cytochrome D ABC transporter ATP binding and permease protein.  
 GN CYCD OR L10706.  
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
 OC NCBI\_Taxid=1360;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=111403;  
 RX MEDLINE=2125186; PubMed=11337471;  
 RA Boletín A., Wincker P., Manger S., Jaillon O., Malarre K.,  
 RA Weissbach J., Ehrlich S.D., Sorokin A.;  
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus  
 RT lactis sep. lactis 111403.";  
 RL Genome Res. 11:731-753(2001).  
 DR EMBL: AB006304; AK04804.1; -;  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR001140; ABC\_TM\_transp.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR Pfam: PF00664; ABC\_membrane\_1.  
 DR Pfam: PF00005; ABC\_tran\_1.  
 DR ProDom: PD000006; ABC\_transporter\_1.  
 DR SMART: SM00382; AAA\_1.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 615 AA; 68037 MW; CD2C5D2BDDBAB22 CRC64;  
 QY Query Match 58.6%; Score 41; DB 16; Length 615;  
 DB Best Local Similarity 66.7%; Pred. No. 85;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 2 THRIHVESAS 10  
 509 THRLHWSS 517  
 RESULT 12

```

08D6G06
ID 08D6G06      PRELIMINARY;      PRT;      863 AA.
AC 08D6G06;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Tl-22259 protein.
GN TlR2259.
OC Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxId=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriiguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsunoto M., Matsuno A., Nakazaki N.,
RA Shimpko S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1."
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005376; BAC09811.1; -.
KW Complete proteome.
SQ SEQUENCE 863 AA; 96927 MW; CEE2CB0B9FF7D01 CRC64;

```

Query Match	58.6%	Score 41	DB 16	Length 863
Best local similarity	45.5%	Pred. No. 1,2e+02		
Matches	5	Conservative	4	Mismatches 2
				Indels 0
				Gaps 0
QY	3	HHHESASLL	13	
		: : : :		
Db	809	HOVHMPVSLI	819	

RESULT 13			
ID	Q8YIY6	PRELIMINARY;	PRT; 229 AA.
AC	Q8YIY6;		
DT	01-MAR-2002 (TREMblrel, 20, Created)		
DT	01-MAR-2002 (TREMblrel, 20, Last sequence update)		
DT	01-MAR-2002 (TREMblrel, 20, Last annotation update)		
DE	Hypothetical cytosolic protein BMEI0303.		
GN	BMEI0303.		
OS	<i>Brucella melitensis</i> .		
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;		
OC	Brucellaceae; Brucella.		
OX	NCBI_TaxID=29459;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=16M / ATCC 23456 / Biotype 1;		
RX	MEDLINE=20020109; PubMed=11756688;		
RA	DeJecchio V.G., Kapatral V., Redkar R.J., Patra G., Mijer C., Los T.,		
RA	Ivanova N., Anderson I., Bhattacharya A., Iykidis A., Reznik G.,		
RA	Rajabonaki L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldman E.,		
RA	Selkov E., Elzer P.H., Hagijs S., O'Callaghan D., Letesson J.-J.,		
RA	HaeslKorn R., Kyriides N., Overbeek R.;		
RT	"The genome sequence of the facultative intracellular pathogen		
RT	<i>Brucella melitensis</i> ."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).		
DR	EMBL; AE009472; AAU51484.1; -		
DR	Hypothetical protein; Complete proteome.		
SQ	SEQUENCE 229 AA; 25507 MW; 98A1769A370F52CA CRC64;		

Query Match	57.1%	Score 40;	DB 16;	Length 229;
Best Local Similarity	58.3%	Pred. No. 45;		
Matches 7; Conservative		2; Mismatches	3; Indels	0; Gaps 0;

QY	1	ITHRIHWESASTL	12
		:         :	
Db	137	IRNRTHWNSANTL	148

```

RESULT 14
ID Q8FYX0 PRELIMINARY; PRT; 229 AA.
AC Q8FYX0;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN BR137.
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; Pubmed=12271122;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
RA Dargatzetis S.C., Debroy R.T., Durkin A.S., Kolonay J.F., Medupu R.,
RA Nelson M.C., Ayodeji B., Kralj M., Shetty J., Malek J., Van Aken S.E.,
RA Rietmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
RT "the Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
RL EMBL; AE014465; AAN30636.1; -.
DR TIGR; BR137; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 229 AA; 25507 MM; 98A1769A370F52CA CRC64;

```

Query Match	Similarity	Score	DB	Length	229;
Best Local	Similarity	58.3%	Pred. No.	45;	
Matches	7;	Conservative	2;	Mismatches	3;
				Indels	0;
				Gaps	0;
Qy	1	ITHRIHWESAST	12		
Db	137	INNRTHWNSANT	148		

```

RESULT 15
ID Q9D912 PRELIMINARY; PRT; 274 AA.
AC Q9D912;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE 1810012L18Rik protein.
GN 1810012L18RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai Y., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Komio H., Adachi J., Fukuda S.,
RA Aizawa K., Iwawa M., Nishi K., Kiyosawa H., Kondo S., Yamamata I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Koehne H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carlinici P., de Bonaldo M.F.,
RA Brownstein M.J., Buit C., Fletcher C., Fujita M., Gariboldi M.,
RA Guernichon S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima T., Mazzarelli J., Mombaerts P.,
RA Noriote P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,

```

RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kontsuki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection.",  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK007461; BAB25050.1; -  
 DR MGD: MG11923740; 181001218rik.  
 DR InterPro: IPR001052; Rubredoxin.  
 DR SMART: SM00355; Znf\_C2H2\_1.  
 DR PROSITE: PS00202; RUBREDOXIN; 1.  
 SQ SEQUENCE 274 AA; 30610 MW; 4BC793D9388E47A9 CRC64;

Query Match 57.1%; Score 40; DB 11; Length 274;  
 Best Local Similarity 85.7%; Pred. No. 55;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 RHIMESA 10  
 DB 76 RHIMESA 82

RESULT 16  
 ID Q18611 PRELIMINARY; PRT; 285 AA.  
 AC Q18611;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical 32.1 kDa protein.  
 GN C44C1.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peleoderinae; Caenorhabditis.  
 NC NCB1\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.",  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Bradshaw H., Steilyes L.,  
 RT "The sequence of C. elegans cosmid C44C1.",  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.,  
 RT "Direct Submision.",  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U41030; AAA82366.2; -  
 DR WormPep: C44C1.1; CE27851.  
 KW Hypothetical protein.  
 SQ SEQUENCE 285 AA; 32140 MW; 3119PDD3CB1212EB CRC64;

Query Match 57.1%; Score 40; DB 5; Length 285;  
 Best Local Similarity 75.0%; Pred. No. 57;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 THRIMES 9  
 DB 211 THRIMES 218

RESULT 17  
 ID Q8PF47 PRELIMINARY; PRT; 338 AA.  
 AC Q8PF47;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical protein XAC4139.  
 GN XAC4139.  
 OS Xanthomonas axonopodis (pv. citri).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Xanthomonas.  
 NC NCB1\_TaxID=92829;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=306 / ATCC 13902 / XV 101;  
 RX MEDLINE=22022145; PubMed=12024217;  
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
 RA Quagho R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
 RA Camarotte G., Cannavan F., Cardoso J., Chambergo F., Ciapina L.P.,  
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Seta J.A.D., Silva C., de Souza R.F.,  
 RA Spinoia L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
 RA Setubal J.C., Kitajima J.P.;  
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities.",  
 RL Nature 417:459-463(2002).  
 DR EMBL: AE012064; AM38974.1; -  
 DR InterPro: IPR001005; Wyp\_DNA\_binding.  
 DR PROSITE: PS00334; MYB\_2; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 338 AA; 37799 MW; 7B9A57D9C01C60BD CRC64;

Query Match 57.1%; Score 40; DB 16; Length 338;  
 Best Local Similarity 58.3%; Pred. No. 68;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 THRIMESASL 13  
 DB 167 THRIOWYCASV 178

RESULT 18  
 ID Q9VR24 PRELIMINARY; PRT; 347 AA.  
 AC Q9VR24;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE CG14821 protein (GH12583p).  
 GN CG14821.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NC NCB1\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Burton R.C., Rogers Y.-H.C., Blazer R.G., Champe O., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abtil J.F., Agbayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballwe R.M., Baer A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Maye A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibeagwam C.,  
 RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mlehnina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murthy L., Muzz D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuseken D.R., Pacleb J.M.,  
 RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RA Ceiniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Batzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresner D., Fartan D.,  
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibeagwam C., Jatali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Piltman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RA Miara S., Crosby M.A., Matthews B.B., Bayraktaroglu I., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Ceiniker S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Kronmiller B., Marshall B., Milburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,  
 RA Asuburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RT "Annotation of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RA Adams M.D., Ceiniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley;  
 RA Stapleton M., Brookslein P., Hong L., Agbayan A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresner D., Fartan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Ceiniker S.;

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE003560; AAF50636.2; -;  
 DR EMBL: AY113342; AAM29347.1; -;  
 DR FlyBase: FBgn005719; CG14821.  
 SQ SEQUENCE 347 AA; 38140 MW; 9FCG45340BE67377 CRC64;  
 Query Match 57.1%; Score 40; DB 5; Length 347;  
 Best Local Similarity 77.8%; Pred. No. 70;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 RHWEASL 12  
 DB 31 RHWESASI 39  
 RESULT 19  
 ID Q9A017 PRELIMINARY; PRT; 541 AA.  
 AC Q9A017;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Hypothetical P protein.  
 GN SPY0975.  
 OS Streptococcus pyogenes.  
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=SP370 / ATCC 700294 / Serotype M1;  
 RX MEDLINE=21192684; PubMed=11296296;  
 RA Ferreretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,  
 RA Pirmieux C., Szatze S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,  
 RA Qian Y., Jia H.G., Najer F.Z., Ren Q., Zhu H., Song L., White J.,  
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
 RT "Complete genome sequence of an M1 strain of *Streptococcus pyogenes*.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).  
 DR EMBL: AB006544; AAK33880.1; -;  
 DR InterPro: IPR006528; PhagesPPI\_SP7.  
 DR TIGRFAMs: TIGR01641; PhagesPPI\_SP7; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 541 AA; 62314 MW; 5F3DF575BF705BC6 CRC64;  
 Query Match 57.1%; Score 40; DB 16; Length 541;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ITHRIHWEA 10  
 DB 446 LTHKGMWDTA 455  
 RESULT 20  
 ID Q8POP2 PRELIMINARY; PRT; 541 AA.  
 AC Q8POP2;  
 DT 01-OCT-2002 (TREMblrel. 22, Created)  
 DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Hypothetical phase protein spyM18\_1275.  
 GN SPYM18\_1275.  
 OS Streptococcus pyogenes (serotype M18).  
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=186103;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=MGAS8232 / Serotype M18;  
 RX MEDLINE=21927593; PubMed=11917108;  
 RA Smoot U.C., Barbican K.D., Van Gompel U.J., Smoot L.M., Chaussee M.S.,  
 RA Sylva G.L., Studevant D.E., Ricklets S.M., Porcella S.F.,  
 RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,

RA Kapur V., Daly J.A., Veasy L.G., Mueser J.M.;  
 RT "Genome sequence and comparative microarray analysis of serotype M18  
 RT group A Streptococcus strains associated with acute rheumatic fever  
 RT outbreak.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).  
 DR EMBL; AE010050; AL97881.1; -;  
 DR InterPro; IPR006528; Phagesp1\_gp7; 1.  
 DR Triframes; TIGR01641; Phagesp1\_gp7; 1.  
 KM Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 541 AA; 62419 MW; F6A0DD87C6939473 CRC64;

Query Match 57.1%; Score 40; DB 16; Length 541;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ITRHWSA 10  
 Db 446 LTRHWSA 455

RESULT 21  
 Q9DBV3 PRELIMINARY; PRT; 1145 AA.  
 AC Q9DBV3; (TREMBlrel. 17, Created).  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE 1200013807RIK protein.  
 GN 1200013807RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=1085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Bonfelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker G., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK004733; BAB23515.1; -;  
 DR MGD; MGI:1918973; 1200013807RIK.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR InterPro; IPR001052; Nubredoxin.  
 DR InterPro; IPR007087; Znf\_C2H2.  
 DR Pfam; PF00271; helicase\_C\_1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC\_C\_1.  
 DR SMART; SM00355; Znf\_C2H2; 1.  
 DR PROSITE; PS00202; RUBREDOXIN; 1.  
 KM ATP-binding; Helicase; Hydrolase.  
 SQ SEQUENCE 1145 AA; 128624 MW; A7B1CDBD77FED4C0 CRC64;

Query Match 57.1%; Score 40; DB 11; Length 1145;  
 Best Local Similarity 85.7%; Pred. No. 2.4e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RHWESA 10  
 Db 947 RHWESA 953

RESULT 22

O8PR28 PRELIMINARY; PRT; 172 AA.

DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Hypothetical protein XAC0138.  
 GN XAC0138.  
 OS Xanthomonas axonopodis (pv. citri).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Xanthomonas.  
 OX NCBI\_TaxID=92829;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=306 / ATCC 13902 / XV 101;  
 RX MEDLINE=22022145; PubMed=12024217;  
 RA da Silva A.C.R., Perro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
 RA Camarotte G., Cannavan F., Cardoso J., Chambergo F., Clapina L.P.,  
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Perro M.I.T.,  
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
 RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
 RA Setubal J.C., Kitajima J.P.;  
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities.";  
 RL Nature 417:459-463(2002).  
 DR EMBL; AE011637; AAM35030.1; -;  
 KM Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 172 AA; 18689 MW; 5A62A85829AD9718 CRC64;

Query Match 55.7%; Score 39; DB 16; Length 172;  
 Best Local Similarity 45.5%; Pred. No. 50;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 3 RHWESASL 13  
 Db 132 H0VHWPMSV 142

RESULT 23

O9HK18 PRELIMINARY; PRT; 197 AA.

DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Endonuclease III related protein.  
 GN TA0790.  
 OS Thermoplasma acidophilum.  
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;  
 OC Xanthomonadaceae; Thermoplasma.  
 OX NCBI\_TaxID=2303;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 1728;  
 RX MEDLINE=20479972; PubMed=11029001;  
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,

RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;  
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma  
acidophilum";  
RL Nature 407:508-513(2000).  
DR EMBL; AL445065; CAC11921.1; -.  
DR HSSP; P20625; ZABK.  
DR InterPro; IPR004035; EndoIII\_FCL.  
DR InterPro; IPR003265; Endo\_3c.  
DR InterPro; IPR003651; Pss\_Bind.  
DR InterPro; IPR000445; HHH\_1.  
DR InterPro; IPR003583; HHH\_1.  
DR Pfam; PF00633; HHH; 1.  
DR Pfam; PF00730; HHH-GPD; 1.  
DR SMART; SM00478; ENDO3c; 1.  
DR SMART; SM00525; Pss; 1.  
DR SMART; SM00278; Hhh1; 1.  
DR PROSITE; PS00764; ENDONUCLEASE\_III\_1; 1.  
KW Complete proteome.  
SQ SEQUENCE 197 AA; 22753 MW; E00E2162AD95A856 CRC64;  
  
Query Match 55.7%; Score 39; DB 17; Length 197;  
Best Local Similarity 70.0%; Pred. No. 58;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 ITHRIHWEASL 10  
Db 126 ISHRIWMSA 135  
  
RESULT 24  
Q92KK1 PRELIMINARY; PRT; 232 AA.  
ID Q92KK1  
AC Q92KK1;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
GN R03328 OR SMC04094.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21396507; PubMed=11481430;  
RA Capela D., Barloy-Hudler F., Gouzy J., Bothe G., Ampe F., Battut J.,  
RA Boistard P., Becker A., Boutry M., Cadieu B., Dreano S., Gloux S.,  
RA Goulet T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,  
RA Poll T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,  
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;  
RT "Analysis of the chromosome sequence of the legume symbiont  
Sinorhizobium meliloti strain 1021.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
DR EMBL; AL591793; CAC47907.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 232 AA; 25861 MW; C73E5DB05764FD67 CRC64;  
  
Query Match 55.7%; Score 39; DB 16; Length 232;  
Best Local Similarity 58.3%; Pred. No. 68;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
  
Qy 1 ITHRIHWEASL 12  
Db 137 IRETHWNSAND 148  
  
RESULT 25  
Q91ZC0 PRELIMINARY; PRT; 321 AA.  
ID Q91ZC0  
AC Q91ZC0;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE G protein-coupled receptor.  
GN MRGB4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6;  
RX MEDLINE=21435808; PubMed=11551509;  
RA Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.;  
RT "A Diverse Family of GPCRs Expressed in Specific Subsets of  
Nociceptive Somatosensory Neurons.";  
RL Cell 106:619-632(2001).  
DR EMBL; AY042202; AAK91798.1; -.  
DR InterPro; IPR000276; GPCR\_Rhodopsin.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PROSITE; PS00237; G\_PROTEIN\_RECER\_F1\_1; 1.  
DR PROSITE; PS50262; G\_PROTEIN\_RECER\_F1\_2; 1.  
KW Receptor.  
SQ SEQUENCE 321 AA; 36337 MW; CBDF185B470DB8B4 CRC64;  
  
Query Match 55.7%; Score 39; DB 11; Length 321;  
Best Local Similarity 54.5%; Pred. No. 96;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 3 HRIHWEASLL 13  
Db 283 HRLQWQSLKLL 293  
  
RESULT 26  
Q91ZB9 PRELIMINARY; PRT; 322 AA.  
ID Q91ZB9  
AC Q91ZB9;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
GN MRGB5.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6;  
RX MEDLINE=21435808; PubMed=11551509;  
RA Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.;  
RT "A Diverse Family of GPCRs Expressed in Specific Subsets of  
Nociceptive Somatosensory Neurons.";  
RL Cell 106:619-632(2001).  
DR EMBL; AY042203; AAK91799.1; -.  
DR InterPro; IPR000276; GPCR\_Rhodopsin.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PROSITE; PS00237; G\_PROTEIN\_RECER\_F1\_1; 1.  
DR PROSITE; PS50262; G\_PROTEIN\_RECER\_F1\_2; 1.  
KW Receptor.  
SQ SEQUENCE 322 AA; 36285 MW; 24E83CBBE0289BBD CRC64;  
  
Query Match 55.7%; Score 39; DB 11; Length 322;  
Best Local Similarity 54.5%; Pred. No. 96;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 3 HRIHWEASLL 13  
Db 284 HRLQWQSLKLL 294  
  
RESULT 27  
Q8RTQ7 PRELIMINARY; PRT; 381 AA.  
ID Q8RTQ7

AC Q8RT07;  
 DT 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
 DE Dissimilatory sulfite reductase alpha subunit (Fragment).  
 GN DSEA.  
 OS Thermodesulfobacterium commune.  
 OC Bacteria; Thermodesulfobacteriales; Thermodesulfobacteriales;  
 OC Thermodesulfobacteriaceae; Thermodesulfobacterium.  
 OX NCBI\_TaxID=1741;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA Klein M., Abicht H.;  
 RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF271771; AAL78311.1; -  
 DR InterPro; IPR006067; Nlr\_Slr\_4Fe4S.  
 DR Pfam; PF01077; Nlr\_Slr\_1.  
 FT NON\_TER 1  
 SQ SEQUENCE 381 AA; 43722 MW; F8C621C35A066712 CRC64;  
 Query Match 55.7%; Score 39; DB 2; Length 381;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

## RESULT 28

ID Q93EV7 PRELIMINARY; PRT; 382 AA.  
 AC Q93EV7;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
 DE Dissimilatory sulfite reductase alpha subunit (Fragment).  
 GN DSEA.  
 OS Thermodesulfobacterium commune.  
 OC Bacteria; Thermodesulfobacteriales; Thermodesulfobacteriales;  
 OC Thermodesulfobacteriaceae; Thermodesulfobacterium.  
 OX NCBI\_TaxID=1741;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA STRAIN=DSM 2178;  
 RC MEDLINE=2145082; PubMed=11567003;  
 RA Klein M., Friedrich M., Roger A.J., Hugenholz P., Fishbain S.,  
 RA Abicht H., Blackall L., Stahl D.A., Wagner M.;  
 RT "Multiple lateral transfers of dissimilatory sulfite reductase genes  
 RT between major lineages of sulfate-reducing prokaryotes";  
 RL J. Bacteriol. 183:6028-6035 (2001).  
 DR EMBL; AF34536; AAK3207.1; -  
 DR InterPro; IPR006067; Nlr\_Slr\_4Fe4S.  
 DR Pfam; PF01077; Nlr\_Slr\_1.  
 FT NON\_TER 1  
 SQ SEQUENCE 382 AA; 43766 MW; 1BBCE4D9398823C8 CRC64;

Query Match 55.7%; Score 39; DB 2; Length 382;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

## RESULT 29

ID Q8CUT5 PRELIMINARY; PRT; 387 AA.  
 AC Q8CUT5;  
 DT 01-MAR-2003 (TReMBLrel. 23, Created)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)

DE Conserved hypothetical protein.  
 GN SC04494 OR SCD35.01 OR SCD69.14.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA STRAIN=A3(2) / M145;  
 RC MEDLINE=21966410; PubMed=12000953;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Klier H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitz E., Rajadream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2)".  
 RL Nature 417:141-147(2002).  
 DR EMBL; AL393120; CAD55486.1; -  
 DR Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 387 AA; 44420 MW; D842730B90A97C4F CRC64;

Query Match 55.7%; Score 39; DB 16; Length 387;  
 Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

## RESULT 30

ID P95619 PRELIMINARY; PRT; 406 AA.  
 AC P95619;  
 DT 01-MAY-1997 (TReMBLrel. 03, Created)  
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE Hydroxyneurosporene synthase.  
 GN CRIC.  
 OS Rhodocyclus gelatinosus (Rhodopseudomonas gelatinosa).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Comamonadaceae; Rhodrivivax.  
 OX NCBI\_TaxID=28068;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA STRAIN=SI;  
 RC Ouchane S., Picard M., Vernotte C., Reiss-Husson F., Astier C.;  
 RT "Pleiotropic effects of puf interposon mutagenesis on carotenoid  
 RT biosynthesis in Rhodrivivax gelatinosus. A new gene organization in  
 RT purple bacteria";  
 RL J. Biol. Chem. 272:1670-1676(1996).  
 DR EMBL; U73944; AAC44799.1; -  
 SQ SEQUENCE 406 AA; 44592 MW; 0CBDA1B64E549D38 CRC64;

Query Match 55.7%; Score 39; DB 2; Length 406;  
 Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

## RESULT 31

ID Q9JPB4 PRELIMINARY; PRT; 406 AA.  
 AC Q9JPB4;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)

```

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Hydroxyneurosporene dehydrogenase.
GN CIRC.
OS Rhodocyclus gelatinosus (Rhodopseudomonas gelatinosa).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Comamonadaceae; Rubrivivax.
OX NCBI_TaxID=28068;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL144;
RA Nagashima K.V., Shimada K., Matsuura K.;
RT "Phylogenetic analysis of photosynthetic genes of Rhodocyclus
RT gelatinosus: Possibility of horizontal gene transfer in purple
RT bacteria.";
RT Photosyn. Res. 36:185-191(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=IL144;
RX MEDLINE=94132007; PubMed=8300574;
RA Nagashima K.V., Matsuura K., Ohyama S., Shimada K.;
RT "Primary structure and transcription of genes encoding B870 and
RT photosynthetic reaction center apoproteins from Rubrivivax
RT gelatinosus.";
RT J. Biol. Chem. 269:2477-2484(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=IL144;
RA Igarashi N., Shimada K., Matsuura K., Nagashima K.V.;
RT "Photosynthetic gene cluster in purple bacterium, Rubrivivax
RT gelatinosus.";
RT (in) Garab G. (eds.);
RT Photosynthesis:
RT mechanisms and effects (Proceedings of the 11th international congress
RT on photosynthesis), pp.4:2889-2892, Kluwer Academic Publishers,
RT Dordrecht (1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=IL144;
RX MEDLINE=20031519; PubMed=10563807;
RA Menin L., Yoshida M., Jéguinod M., Nagashima K.V., Matsuura K.,
RA Patoc P., Vermeglio A.;
RT "Dark aerobic growth conditions induce the synthesis of a high
RT midpoint potential cytochrome c8 in the photosynthetic bacterium
RT Rubrivivax gelatinosus.";
RT Biochemistry 38:15238-15244(1999).
DR EMBL: AB034704; BAA94047.1;
SQ SEQUENCE 406 AA; 44797 MW; ASAFAF4045859846 CRC64;

Query Match 55.7%; Score 39; DB 2; Length 406;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

CY 4 RHIMESASL 13
DB 192 RLHWDSIL 201

RESULT 32
ID 08FU88 PRELIMINARY; PRT; 410 AA.
AC 08FU88;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Putative membrane protein.
GN CE0133.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikeo K., Suzuki M., Maehama J., Itoh T., Yamagishi A., Nishio Y.,
RA Ueda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";
RT Submitted (May-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AP005214; BAC16943.1;
KV Complete proteome.
SQ SEQUENCE 410 AA; 43418 MW; 07F91D9B5FD6D26A CRC64;

Query Match 55.7%; Score 39; DB 16; Length 410;
Best Local Similarity 46.2%; Pred. No. 1.2e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

CY 1 ITHRHESASL 13
DB 225 LVHGVNKSASAL 237

RESULT 33
ID 08MMP7 PRELIMINARY; PRT; 437 AA.
AC 08MMP7;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Methionyl-tRNA synthetase beta subunit. 6/101.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachepat J., Dear P.,
RA Lehmann R., Baumann C., Parra G., April J.F., Gutgo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostellium.";
RT Submitted (May-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AC117080; AAM45346.1;
DR InterPro: IPR004803; QCRNA_ribo_trans.
DR InterPro: IPR002616; tRNA_Tibo_Trans.
DR Pfam: PF01702; TGT; 1.
DR TIGRFAMs: TIGR00430; Q_CRNA_tgc; 1.
DR TIGRFAMs: TIGR00449; tgc_general; 1.
SQ SEQUENCE 437 AA; 49833 MW; 7AF7555D99427A6 CRC64;

Query Match 55.7%; Score 39; DB 5; Length 437;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

CY 2 THRHESASL 13
DB 243 TH-HWERTSL 252

RESULT 34
ID 08B3H4 PRELIMINARY; PRT; 572 AA.
AC 08B3H4;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN GB51785.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,

```



```

RA Masdek T., Zouine M., Couve E., Lallouf L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513 (2002).
DR EMBL; AL766853; CAD47444.1; -.
DR Sagaliet; gba1785; -.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 572 AA; 64458 MW; 9CB68B07FA063F9E CRC64;

Query Match 55.7%; Score 39; DB 16; Length 572;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 2 THRIHW 7
Db 524 THRIHW 529

RESULT 35
Q8DXV6 PRELIMINARY; PRT; 572 AA.
AC Q8DXV6;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE ABC transporter, ATP-binding protein CydD.
GN SAG1740.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RA SEQUENCE FROM N.A.
RA STRAIN=2603 V/R / Serotype V;
RA MEDLINE=22222988; PubMed=12200547;
RA Tetteijn H., Maesigant V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Weesje M.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beaman M.J., Brinkac L.M., Daugherty S.C.,
RA Deacy R.T., Durkin A.S., Kolman J.F., Madupu R., Lewis M.R.,
RA Radue D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.B., Gill J., Scarselli M., Mora M.,
RA Jacobini E.T., Brettoni C., Gall G., Mariani M., Vegni P., Malone D.,
RA Rinaldo D., Rappaport R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).
DR EMBL; AE014270; AAN00603.1; -.
DR TIGR; SAG1740; -.
KM ATP-binding; Complete proteome.
SQ SEQUENCE 572 AA; 64423 MW; C3B2FC28C3D47A3F CRC64;

Query Match 55.7%; Score 39; DB 16; Length 572;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 2 THRIHW 7
Db 524 THRIHW 529

RESULT 36
Q927C5 PRELIMINARY; PRT; 574 AA.
AC Q927C5;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE CydC protein.
GN CydC OR LIN2864.
OS Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

```

```

OX NCBI_TaxID=1642;
RN [1]
RA SEQUENCE FROM N.A.
RA STRAIN=CLIP 11262 / Serovar 6a;
RC PubMed=11679669;
RX Glaser P., Frangeul L., Buchrieser C., Ruenliok C., Amend A.,
RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA Chablit A., Checouani F., Couve E., de Darvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Duseurget O.,
RA Entian K.-D., Feihl H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krefl J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Matcounam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Coessart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852 (2001).
DR EMBL; AL596173; CAC98090.1; -.
DR ListList; LIN02864; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
KM Complete proteome.
SQ SEQUENCE 574 AA; 64042 MW; B5B3A9758CAF8B98 CRC64;

Query Match 55.7%; Score 39; DB 16; Length 574;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 2 THRIHW 7
Db 527 THRIHW 532

RESULT 37
Q8Y3W3 PRELIMINARY; PRT; 574 AA.
AC Q8Y3W3;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE CydC protein.
GN CydC OR LMO2716.
OS Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RA SEQUENCE FROM N.A.
RA STRAIN=EGD-e / Serovar 1/2a;
RC MEDLINE=21537279; PubMed=11679669;
RX Glaser P., Frangeul L., Buchrieser C., Ruenliok C., Amend A.,
RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA Chablit A., Checouani F., Couve E., de Darvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Duseurget O.,
RA Entian K.-D., Feihl H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krefl J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Matcounam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Coessart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852 (2001).
DR EMBL; AL591984; CAB00929.1; -.
DR ListList; LMO02716; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transporter.

```

DR Pfam: PF00664; ABC membrane; 1.  
 DR Pfam: PD00005; ABC trans; 1.  
 DR Pfam: PD00006; ABC transporter; 1.  
 DR SMART: SM00382; AAA; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 574 AA; 64075 MW; 6A9B605E4CCCB63 CRC64;

Query Match 55.7%; Score 39; DB 16; Length 574;  
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 THRIHW 7  
 |||:  
 Db 527 THRIHW 532

## RESULT 38

Q9NS78 PRELIMINARY; PRT; 250 AA.

AC Q9NS78; 01-OCT-2000 (TREMblrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)

DE 01-MAR-2003 (TREMblrel. 23, Last annotation update)

GN SMN interacting protein 1-gamma.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI\_TaxID=9606;

RE [1]

SEQUENCE FROM N.A.

RC TISSUE=spinal cord;

RA Aetajinal W., Arahata K., Tsukahara T.;

RT "Increased expression level of a novel alternative splicing variant of

RT SIPI in motor neuron diseases.";

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB037702; BAB03509.1;

DR InterPro; IPR007022; SIPI.

DR Pfam; PF04938; SIPI; 1.

SQ SEQUENCE 250 AA; 28155 MW; 4DD0F37BD5B7C837 CRC64;

Query Match 54.3%; Score 38; DB 4; Length 250;  
 Best Local Similarity 60.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 HRIHESASL 12  
 |||:  
 Db 120 HRIHESASL 129

## RESULT 39

Q9NS79 PRELIMINARY; PRT; 265 AA.

AC Q9NS79; 01-OCT-2000 (TREMblrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)

DE 01-MAR-2003 (TREMblrel. 23, Last annotation update)

GN SMN interacting protein-beta.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI\_TaxID=9606;

RE [1]

SEQUENCE FROM N.A.

RC TISSUE=spinal cord;

RA Aetajinal W., Arahata K., Tsukahara T.;

RT "Increased expression level of a novel alternative splicing variant of

RT SIPI in motor neuron diseases.";

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB037701; BAB03508.1;

DR InterPro; IPR007022; SIPI.

DR Pfam; PF04938; SIPI; 1.

SQ SEQUENCE 265 AA; 29931 MW; FE91C496F663B825 CRC64;

Query Match 54.3%; Score 38; DB 4; Length 265;  
 Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 HRIHESASL 12  
 |||:  
 Db 120 HRIHESASL 129

## RESULT 40

Q8UBV4 PRELIMINARY; PRT; 266 AA.

ID Q8UBV4; 01-JUN-2002 (TREMblrel. 21, Created)

DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)

DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)

GN Hypothetical protein Atu2746.

OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

CC Rhizobiaceae; Rhizobium.

NCBI\_TaxID=17629;

RE [1]

SEQUENCE FROM N.A.

RC MEDLINE=21608550; PubMed=11743193;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,

RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Woo L.,

RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyee D. Sr.,

RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,

RA Kutayavin T., Levy R., Li M.-D., McClelland E., Palmeri A.,

RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,

RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,

RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,

RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,

RA Neater E.W.;

RT "The genome of the natural genetic engineer Agrobacterium tumefaciens

RT C58.";

RL Science 294:2317-2323(2001).

RE [2]

SEQUENCE FROM N.A.

RC MEDLINE=21608551; PubMed=11743194;

RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,

RA Houmlel K., Goldman B.S., Cao Y., Akenazi M., Halling L.,

RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,

RA Flanagan C., Crowell C., Gurrison J., Lomo C., Sear C., Strub G.,

RA Cielo C., Slater S.;

RT "Genome sequence of the plant pathogen and biotechnology agent

RT Agrobacterium tumefaciens C58.";

RL Science 294:2323-2328(2001).

DR EMBL; AE009222; AAL43727.1; ALT\_INIT.

DR EMBL; AE008188; AAK8461.1;

SQ SEQUENCE 266 AA; 29620 MW; 40B98CC626C43B27 CRC64;

Query Match 54.3%; Score 38; DB 16; Length 266;  
 Best Local Similarity 58.3%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 ITHRIHESASL 12  
 |||:  
 Db 169 ITHRIHESASL 180

Search completed: August 28, 2003, 14:12:50  
 Job time : 95 secs